Product Differentiation and Oligopoly: a Network Approach

Bruno Pellegrino

University of Maryland

July 27, 2021

Abstract

Industry concentration and corporate profit rates have increased, in the United States, over the past two decades. This paper investigates the welfare implications of economic activity concentrating within a few firms that hold market power. I develop a general equilibrium model that features granular firms that compete in a network game of oligopoly, alongside a competitive fringe of atomistic firms with endogenous entry. To capture the degree of product differentiation among the oligopolists, I introduce a Generalized Hedonic-Linear (GHL) demand system. I show how to identify this demand system using a publicly-available dataset that measures product similarity among all public corporations in the US. Using my model, I estimate a large deadweight loss from oligopolistic behavior, equal to 11% of the total surplus produced by public firms. This loss would increase to 20% if all these firms were allowed to collude. The distributional effects of oligopoly are quantitatively important as well: under perfect competition, consumer surplus would double with respect to the oligopolistic equilibrium. I also estimate that the deadweight loss has increased by at least 2.5 percentage points since 1997. The share of surplus that accrues to producers as profits also has increased. Finally, I show how the dramatic rise in startups' proclivity to sell off to incumbents (rather than go public) may have contributed to these trends.

JEL Codes: D2, D4, D6, E2, L1, O4
Keywords: Competition, Concentration, General Equilibrium, Market Power, Markups, Mergers, Monopoly, Networks, Oligopoly, Startups, Text Analysis, Welfare

This paper received the following awards: EEA Best Job Market Paper Award, WFA Best Corporate Finance Paper, UCLA Xavier Drèze Prize for Outstanding Research Paper and Best Presentation Award at the Unicredit Job Market Bootcamp. I gratefully acknowledge financial support from the Price Center for Entrepreneurship at UCLA Anderson. I am indebted to my dissertation advisors Hugo Hopenhayn, Nico Voigtländer, Andy Atkeson, Ed Leamer, Stavros Panageas and Romain Wacziarg for their invaluable advice and guidance. I benefited from thoughtful conference discussions by Michael Gofman, Naveen Khanna, Isabelle Mejean and Jane Olmstead-Rumsey. I also received valuable comments from John Asker, David Baqaee, Alberto Bisin, Ariel Burstein, Michela Giorcelli, Thomas Groesbeck, Kurt Mitman, Oleg Itskhoki, Nicola Persico, Sebastian Ottinger, Tommaso Valletti and Maria Lucia Yanguas. Finally, I acknowledge outstanding data collection assistance by Michele Schincaglia.

1. Introduction

Industry concentration, markups, and corporate profit rates have all increased in the United States during the past two decades (Grullon et al., 2019; De Loecker et al., 2020). This fact has spurred important public debates over whether these trends reflect an oligopolization of U.S. industries and whether a revised antitrust paradigm is necessary (Khan, 2018). While standard price theory arguments suggest that the welfare implications of these trends might be significant, interpreting these trends presents an imposing methodological challenge. The study of market power has traditionally resided within the literature on Empirical Industrial Organization (EIO). Yet, there is a consensus that these trends are macroeconomic in nature: standard EIO methodologies are unfeasible, as they require data that is not available for more than a handful of industries (Syverson, 2019).

This paper investigates the welfare consequences of the increased concentration of U.S. industries. I address the existing methodological challenges by introducing a novel general equilibrium model with two types of firms: a finite set of *granular* firms that behave as oligopolists and a continuum of atomistic producers that behave competitively and can enter and exit. To model product market competition among the oligopolists, I propose a Generalized Hedonic-Linear (GHL) demand system, which I estimate using a dataset recently developed by Hoberg and Phillips (2016). This dataset provides measures of product similarity for all pairs of publicly-traded corporations in the US. The empirical implementation of the model allows me address the following question: how have consumer surplus and the welfare costs of oligopoly evolved as a consequence of industry consolidation during this period?

Using my novel theoretical framework, I document that the increased concentration of US industries over the past twenty years was accompanied by an increase in oligopoly power, measured as: (1) an increase in the deadweight losses induced by oligopolistic behavior; (2) a decline in the share of total surplus that accrues to consumers. My methodology also allows to me associate these trends with another well-known stylized fact: the dramatic rise in takeovers of startups that began in the mid 1990s, and which coincided with the well-known secular decline in Initial Public Offerings (IPOs) (Kahle and Stulz, 2017).

Economists have long been concerned with market power. Since the 1980s, the EIO literature has been developing a conceptual "toolkit" that researchers and antitrust enforcement practitioners have used to analyze market power within industries (Einav and Levin, 2010). The EIO approach requires the researcher to first understand the structure of product market rivalries in an industry: a firm's ability to price above marginal cost depends critically on the intensity of competition from firms that produce similar products. As a consequence, this literature has shown that oligopoly power is inextricably linked to the notion of product differentiation: to measure a firm's market power in an industry with n firms, the economist effectively needs to first estimate n^2 cross-price demand elasticities—one for each pair of rivals. This is because variations in the supply of any product cause the residual demand curve of every competing product to shift, and the shift is larger for products that are closer substitutes.

In industry studies, demand estimation is nowadays largely based on hedonic models (Berry, Levinsohn and Pakes, 1995). The current resurgence in market power and antitrust research, however, has a distinctive macroeconomic angle. Because we do not observe output volume, prices, or product characteristics for a sufficiently large cross-section of industries, this EIO approach cannot be directly applied in a macroeconomic context. This challenge is compounded by the problem that, at the macro level, product-market rivalry is not well approximated by industry classifications. Industry classifications (such as NAICS) tend to be

based on similarities in the production process, not on the degree of product substitutability. In other words, they are appropriate for estimating production functions, but they are unreliable when it comes to estimating cross-price demand elasticities. In addition, the very concepts of industry or sector are more fluid than macroeconomists tend to assume. While industry classifications are static, larger companies (those more likely to have market power) move frequently from one industry to another, and have been shown to strategically manipulate their industry classification—a phenomenon that has been dubbed *industry window dressing* (Chen et al., 2016).

Despite these challenges, the macroeconomics literature has made significant progress in incorporating market power in general equilibrium models: Baqaee and Farhi (2020, henceforth BF) have recently shown how to approximate the welfare costs of imperfect competition, under minimal assumptions, using the cross-sectional distribution of markups. Their approach is (by design) very general, but it is also agnostic about the origin of markups, which are assumed to be exogenous and observable. Because this approach is silent about how the observed cross-section of markups arises in the first place, a separate theory of markups formation is needed in order to simulate the welfare consequences of a change in market structure.

This study breaks new ground by providing a theory of firm size and profitability that generalizes the Cournot oligopoly model to differentiated products and hedonic demand, and embeds it in a general equilibrium model. The objective of my model is not to capture all sources of variation in markups, but rather to isolate the variation that can be reliably attributed to product market rivalry. Through this approach, I can quantify the contribution of each individual producer to aggregate welfare, and I can study the general equilibrium effects of events that are relevant to antitrust policy, such as mergers or the entry of additional firms.

To achieve this, my theoretical model dispenses with the notions of industry and sector altogether, building instead on the tradition of hedonic demand (Lancaster, 1966; Rosen, 1974). Thus, I can link the cross-price elasticity of demand between all firms in the economy to the fundamental attributes of each firm's product portfolio. Each firm's output is modeled as a bundle of characteristics that are individually valued by the representative consumer. The cross-price elasticity of demand between two firms depends on the characteristics embedded in their output. If the product portfolios of two companies contain similar characteristics, the cross-price elasticity of demand between their products is high. The result is a rather different picture of the product market: not a collection of sectors, but a network, in which the distance between nodes reflects product similarity and strategic interaction between firms.

The key assumptions of my model are: (1) the representative consumer's preferences are described by a utility function that is quadratic-in-characteristics ; (2) firms compete à la Cournot[1]; (3) the marginal cost function is linear in output. Based on these assumptions, the firms in my model play a linear-quadratic game over a weighted network, a type of potential game that has been extensively studied in the micro theory literature (see Ballester, Calvó-Armengol and Zenou, 2006; Ushchev and Zenou, 2018).

This is the first paper to show how to derive the network Cournot model starting from a hedonic utility specification, to embed the game in a general equilibrium framework and to take the model to the data in a structural way.

To estimate my model, I use a recently-developed data set (Hoberg and Phillips, 2016, henceforth HP) that provides continuous measures of product similarity for every pair of publicly traded firms in the United States.

[1]I also study the Bertrand case in the Appendix.

These product-similarity scores are based on a computational-linguistics analysis of product descriptions obtained from SEC filings, and are mapped by my model into an $n \times n$ matrix of cross-price demand elasticities. Moreover, because HP's similarity scores are time-varying (yearly observations since 1997), my model is unique in that the degree of product substitution between individual firms is allowed to change over time.

Crucially, the empirical implementation of my model does not require any proprietary or confidential data, and is computationally tractable. Two datasets are required: Compustat and HP's cosine similarity data, which the authors have made publicly-accessible through an online repository.[2]

I use my model to compute the deadweight loss from oligopoly and to simulate changes in total surplus and consumer surplus for a number of counterfactuals. I find that the welfare costs of oligopoly are sizable. By moving to an allocation in which firms price at marginal cost (that is, in which they behave as if they were atomistic players in a perfectly-competitive market), total surplus would rise by approximately 11 percentage points; consumer surplus would double, partly due to total surplus being reallocated from producers to consumers. By computing a separate counterfactual that keeps the aggregate labor supply fixed (markups are equalized, rather than eliminated), I can determine that a significant share of the welfare loss from oligopoly—about 7.7 percentage points of the aforementioned 11—occurs by way of factor misallocation. In other words, the deadweight loss is driven not only by an underutilization of inputs, but also by a suboptimal mix of goods being produced. I also simulate a counterfactual in which all firms in the economy are owned by a single producer that implements a collusive equilibrium. Under this scenario, total surplus would drop by about one-tenth: with some degree of abstraction, we can think of this estimate as an upper bound to the welfare benefits of antitrust. Also, in this monopolistic/collusive equilibrium consumer surplus would decrease by about 38%, due partly to surplus being reallocated from consumers to producers.

By mapping my model to firm-level data for a period of 21 consecutive years, I investigate the welfare consequences of the rise in concentration and markups between 1997 and 2017. I find that the share of surplus appropriated by companies in the form of oligopoly profits has increased from about 50% (in 1997) to 55.5% (in 2017). When I subtract fixed costs (such as capital and overhead) from profits and total surplus, this increase becomes significantly steeper: from 11% in 1997 to 22% in 2017. This result is robust to different measurements of fixed costs and intangible capital, and suggests that the increase in the profit share of surplus is not justified by larger fixed costs.

The welfare costs of oligopoly have also increased over this period. In terms of total surplus, the gap between the oligopolistic equilibrium and perfect competition (the deadweight loss) has increased from 8.5% (in 1997) to 11% (in 2017). Consumer surplus is adversely affected via two channels: less surplus is produced overall (as a percentage of the surplus that *could* be produced), and less of the diminished surplus is allocated to the consumer in equilibrium. Thus, an important contribution of this paper is to investigate the distributional implications of oligopoly. Overall, my empirical findings suggest that rising concentration and markups did indeed translate into a measurable increase in the welfare loss from oligopoly, and affected how surplus is shared between producers and consumers.

Finally, I use the counterfactual-building capabilities of the model to better understand the drivers of these trends. In particular, I study the impact of a dramatic secular shift that has occurred among venture capital (VC) startups over the past 20 years. In the early 1990s, most VC-backed startups (80%–90%), if

[2]See hobergphillips.tuck.dartmouth.edu

successful, would exit[3] through an Initial Public Offering (IPO). Today, the near entirety (about 94%) of successful VC-backed startups exit by getting acquired by an incumbent. I find that this shift accounts not only for the secular decline in the number of public corporations in the US (from about 7,500 in 1997 to about 3,500 in 2017) but also for a large share of the increase in the welfare costs of oligopoly, and of the rise in the profit share of surplus.

This paper aims to connect the new EIO literature (Einav and Levin, 2010) to two recent and growing branches of macroeconomics that use micro-data.

The first is the literature on networks (Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi, 2012; Carvalho, 2014; Acemoglu, Ozdaglar and Tahbaz-Salehi, 2017; Carvalho and Tahbaz-Salehi, 2019; Baqaee and Farhi, 2020; Carvalho, Nirei, Saito and Tahbaz-Salehi, 2020). I contribute to and expand this literature, which has mostly focused on input-output networks, by considering a different type of network: that of product market rivalries.[4]

The second is the literature on markups and industry concentration (De Loecker, Eeckhout and Unger, 2020; Autor, Dorn, Katz, Patterson and Van Reenen, 2020; Edmond, Midrigan and Xu, 2018; Covarrubias, Gutiérrez and Philippon, 2020; Syverson, 2019). This paper builds on and adds to this body of work by incorporating hedonic demand as well as new data. These features allow me to go beyond markups and concentration, and to create a rich, high-dimensional representation of the competitive environment. In my model, firms differ not only by their productivity, but also by their products' characteristics; as a consequence, each firm has a distinct set of competitors that changes over time, as firms update their product's description in their SEC filings.

This paper also connects the recent literature on market power to the secular decline of public companies (Kahle and Stulz, 2017) and IPOs (Bowen, Frésard and Hoberg, 2018; Gao, Ritter and Zhu, 2013). My model allows to quantify the effects of these phenomena on the intensity of product market competition.

The rest of the paper is organized as follows. In Section 2, I present my theoretical model. In Section 3, I present the data used in the empirical part of the paper and show how it is mapped to the model. In Section 4, I present my empirical results. In Section 5, I discuss a number of extensions and robustness checks. In Section 6, I present my conclusions and discuss how my findings can inform the current debate on market power and antitrust policy.

[3] In the entrepreneurial finance literature, an "exit" is the termination of a VC investment and should not be confused with a business termination. If the VC investor exits with an IPO, that event marks the entry of that firm in the universe of public firms, not an enterprise death.

[4] Outside the macro-networks literature, Bloom, Schankerman and Van Reenen (2013) have studied rivalry networks in a seminal empirical study of R&D spillovers.

2. A Theory of Imperfect, Networked Competition

In this section, I present a general equilibrium model in which firms produce differentiated products and compete à la Cournot. For expositional purposes, I start by laying out the basic model that only includes granular oligopolistic firms. After characterizing the equilibrium of this model economy and outlining a series of counterfactuals of interest, I extend the model (in Subsection 2.6) by adding a continuum of perfectly-competitive atomistic firms.

2.1. Basic Setup: the Generalized Hedonic-Linear (GHL) Demand System

There are n firms, indexed by $i \in \{1, 2, ..., n\}$ that produce differentiated products. Following the tradition of hedonic demand in differentiated product markets (Lancaster, 1966; Rosen, 1974), I assume that consumers value each product as a bundle of characteristics. The number of characteristics is $k + n$.

There are two types of characteristics. The first k characteristics are common across all goods and are indexed by $j \in \{1, 2, ..., k\}$, while the remaining n characteristics are idiosyncratic (that is, they are product-specific and cannot be imitated by other products) and therefore have the same index i as the corresponding product. The scalar a_{ji} is the number of units of common characteristic j provided by product i. Each product is described by a k-dimensional column vector $\mathbf{a}_i$, which I assume (without loss of generality) to be of unit length – formally:

$$\mathbf{a}_i = \begin{bmatrix} a_{1i} & a_{2i} & \dots & a_{ki} \end{bmatrix}' \tag{2.1}$$

$$\text{such that} \qquad \sum_{j=1}^{k} a_{ji}^2 = 1 \quad \forall\, i \in \{1, 2, ..., n\} \tag{2.2}$$

The vector $\mathbf{a}_i$ therefore provides firm i's coordinates in the space of common characteristics. We can stack all the coordinate vectors $\mathbf{a}_i$ inside a $k \times n$ matrix that we call $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix} \tag{2.3}$$

Let q_i be the number of units produced by firm i and consumed by the representative agent, which we write inside the n-dimensional vector $\mathbf{q}$:

$$\mathbf{q} = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}' \tag{2.4}$$

Definition 1. A vector $\mathbf{q}$ that specifies, for every firm, the number of units produced is called an *allocation.*

I assume that there exists a representative agent. Consistent with the hedonic demand literature, the consumer combines linearly the characteristics of different products, and their preferences are defined in terms of these characteristics. Letting x_j be the total units of common characteristic j, we have:

$$x_j = \sum_i a_{ji} q_i \tag{2.5}$$

Hence, geometrically, the matrix $\mathbf{A}$ projects the vector of units purchased $\mathbf{q}$ on the space of common characteristics:

$$\mathbf{x} \;=\; \mathbf{A}\mathbf{q} \tag{2.6}$$

With regard to the n idiosyncratic characteristics, I assume that each unit of good i provides exactly one unit of its corresponding idiosyncratic characteristic. Hence, we can just write q_i in place the units of idiosyncratic characteristic i.

The representative agent's preferences are described by a utility function that is quadratic in both common characteristics ($\mathbf{x}$) and idiosyncratic characteristics ($\mathbf{q}$). The agent's preferences also incorporate a linear disutility for the total number of hours of work supplied (H):

$$U(\mathbf{x}, \mathbf{q}, H) \;\stackrel{\text{def}}{=}\; \alpha \cdot \sum_{j=1}^{k} \left(b_j^x x_j \,-\, \frac{1}{2}\, x_j^2 \right) + (1-\alpha) \sum_{i=1}^{n} \left(b_i^q q_i \,-\, \frac{1}{2}\, q_i^2 \right) - H \tag{2.7}$$

where b_j^x and b_i^q are characteristic-specific preference shifters. In linear algebra notation:

$$U(\mathbf{x}, \mathbf{q}, H) \;\stackrel{\text{def}}{=}\; \alpha \left(\mathbf{x}'\mathbf{b}^x - \frac{1}{2} \cdot \mathbf{x}'\mathbf{x} \right) + (1-\alpha) \left(\mathbf{q}'\mathbf{b}^q - \frac{1}{2} \cdot \mathbf{q}'\mathbf{q} \right) - H \tag{2.8}$$

$\alpha \in [0,1]$ is the utility weight that is assigned to common characteristics. Hence, it governs the degree of *horizontal differentiation* among products. This utility specification is a generalization of the preferences used by Epple (1987). In addition to introducing idiosyncratic characteristics, I make leisure the outside good: that allows me to close the model and make it general equilibrium.

I denote by h_i the labor input acquired by every firm, so that the labor market clearing condition is:

$$H \;=\; \sum_i h_i \tag{2.9}$$

I assume (without loss of generality) that labor is the numéraire of this economy (the price of one unit of labor is 1\$), therefore h_i is also the total variable cost incurred by firm i. Firm i produces output q_i using a quasi-Cobb Douglas production function:

$$q_i \;=\; k_i^{\theta} \cdot \ell(h_i) \tag{2.10}$$

where k_i is the capital input (fixed) and the function $\ell(\cdot)$ is such that firm i's technology can be described by the following quadratic total variable cost function:

$$h_i \;=\; c_i q_i \,+\, \frac{\delta_i}{2}\, q_i^2 \tag{2.11}$$

where c_i and δ_i depend on k_i. MC and AVC denote, respectively, the marginal cost and the average variable cost:

$$\mathrm{MC}_i \;=\; c_i + \delta_i q_i; \qquad \mathrm{AVC}_i \;=\; c_i + \frac{\delta_i}{2} q_i \tag{2.12}$$

For some of the empirical analysis, I will later also consider fixed costs (f_i). Firm i's total cost function will then become:

$$\mathrm{TC}_i \;=\; f_i + c_i q_i + \frac{\delta_i}{2} q_i^2 \tag{2.13}$$

The representative consumer buys the goods bundle $\mathbf{q}$ taking $\mathbf{p}$ (the vector of prices) as given. Moreover,

I assume that the representative consumer is endowed with the shares of all the companies in the economy. As a consequence, the aggregate profits are paid back to them. Their consumption basket, defined in terms of the unit purchased $\mathbf{q}$, has to respect the following budget constraint:

$$H + \Pi \;=\; \sum_{i=1}^{k} p_i q_i \tag{2.14}$$

Notice that for now we have defined aggregate economic profits Π to include all non-labor compensation (which equates to assuming that f_i is sunk). We will later consider a narrower metric of profits from which fixed costs $\left(F \overset{\text{def}}{=} \sum_i f_i\right)$ are netted out.

2.2. Equilibrium

To streamline notation, let us define:

$$b_i \quad \overset{\text{def}}{=} \quad \alpha \sum_j a^x_{ji} x_j + (1-\alpha)\, b^q_i \tag{2.15}$$

or, in linear algebra notation:

$$\mathbf{b} \quad \overset{\text{def}}{=} \quad \alpha\, \mathbf{A}'\mathbf{b}^x + (1-\alpha)\,\mathbf{b}^q \tag{2.16}$$

Then, plugging equation (2.6) and (2.16) inside equation (2.8), we obtain the following Lagrangian for the representative consumer:

$$\mathscr{L}(\mathbf{q}, H) \quad = \quad \mathbf{q}'\mathbf{b} - \frac{1}{2}\mathbf{q}'\left[\mathbf{I} + \alpha\left(\mathbf{A}'\mathbf{A} - \mathbf{I}\right)\right]\mathbf{q} - H - \lambda\left(\mathbf{q}'\mathbf{p} - H - \Pi\right) \tag{2.17}$$

The choice of labor hours as the numéraire immediately pins down the Lagrange multiplier $\lambda = 1$. Then, the consumer chooses a demand function $\mathbf{q}(\mathbf{p})$ to maximize the following consumer surplus function:

$$S(\mathbf{q}) \quad = \quad \mathbf{q}'(\mathbf{b} - \mathbf{p}) - \frac{1}{2}\mathbf{q}'\left[\mathbf{I} + \alpha\left(\mathbf{A}'\mathbf{A} - \mathbf{I}\right)\right]\mathbf{q} \tag{2.18}$$

Let us now define the concept of *cosine similarity.*

Definition 2. We call the dot product $\mathbf{a}_i'\mathbf{a}_j$ the *cosine similarity* between i and j.

The rationale for this nomenclature is that – geometrically – $\mathbf{a}_i'\mathbf{a}_j$ measures the cosine of the angle between vectors $\mathbf{a}_i$ and $\mathbf{a}_j$ in the space of common characteristics $\mathbb{R}^k$. Hence, the cosine similarity ranges from zero to one. Because, by definition:

$$\left(\mathbf{A}'\mathbf{A}\right)_{ij} \;=\; \mathbf{a}_i'\mathbf{a}_j \tag{2.19}$$

the matrix $\mathbf{A}'\mathbf{A}$ contains the *cosine similarities* between all firm pairs. A higher cosine similarity implies that two products provide a more overlapping mix of characteristics, and this reflects in patterns of product substitution: if $\mathbf{a}_i'\mathbf{a}_j > \mathbf{a}_i'\mathbf{a}_k$, an increase in the supply of product i leads to a larger decline in the marginal utility of product j than it does on the marginal utility of product k.

Figure 1 helps visualize this setup for the simple case of two firms—1 and 2—competing in the space of two common characteristics A and B. As can be seen in the figure, both firms exist as vectors on the unit

Figure 1: Example Product Space: Two Firms, Two Characteristics

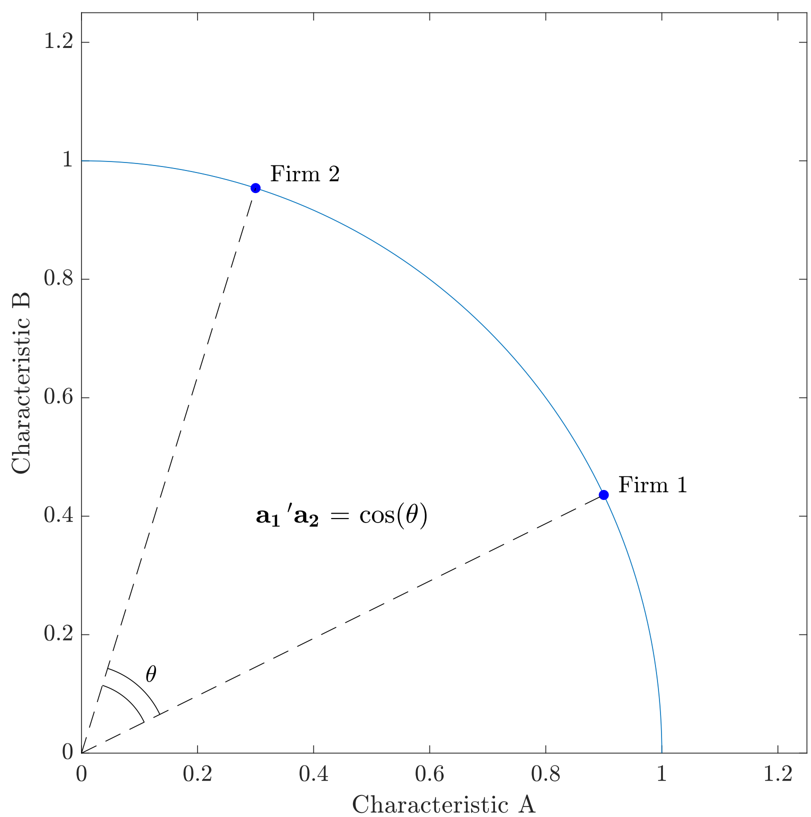


Figure Notes: The following diagram exemplifies the hedonic demand model, for the simple case where there are only two product characteristics (A and B) and only two competitors (1 and 2). Each firm exists as a vector on the unit hypersphere of product characteristics (in this example, we have a circle). The dot product $\mathbf{a}_i'\mathbf{a}_j$ equals the cosine of the angle θ. The tighter the angle, the higher the cosine similarity, and the larger (in absolute value) the inverse cross-price elasticity of demand.

circle (with more than three characteristics, it would be a hypersphere instead). The cosine similarity $\mathbf{a}_i'\mathbf{a}_j$ captures the width of the angle θ. An increase in the cosine of the angle θ implies a lower angular distance, and therefore a more overlapping set of common characteristics.

The assumption that $\mathbf{a}_i$ has unit length is a normalization assumption on volumetric units (kilograms, pounds, gallons, etc.). The normalization consists in picking, for each good i, the volume unit so that i is geometrically represented by a point on the k-dimensional hypersphere.

We can streamline the notation further by defining:

$$\mathbf{\Sigma} \stackrel{\text{def}}{=} \alpha\left(\mathbf{A}'\mathbf{A} - \mathbf{I}\right) \tag{2.20}$$

then the demand and inverse demand functions are given by:

$$\text{Aggregate demand}: \quad \mathbf{q} \; = \; (\mathbf{I}+\mathbf{\Sigma})^{-1}(\mathbf{b}-\mathbf{p}) \tag{2.21}$$

$$\text{Inverse demand}: \quad \mathbf{p} \; = \; \mathbf{b}-(\mathbf{I}+\mathbf{\Sigma})\,\mathbf{q} \tag{2.22}$$

Notice that the quantity sold by each firm may affect the price of the output sold by every other firm in the economy (unless the matrix $\mathbf{\Sigma}$ is null). The derivative $\partial p_i/\partial q_j$ is proportional to $\mathbf{a}_i'\mathbf{a}_j$, the product similarity between i and j. The closer these two firms are in the product characteristics space, the larger is this derivative in absolute value. Because $\mathbf{A}'\mathbf{A}$ is symmetric, we have $\partial q_i/\partial p_j = \partial q_j/\partial p_i$ by construction. My rationale for using a linear demand is discussed at length in Appendix G.

In terms of elasticities, we have:

$$\text{Inverse cross}-\text{price elasticity of demand}: \quad \frac{\partial \log p_i}{\partial \log q_j} \; = \; -\frac{q_j}{p_i}\cdot\sigma_{ij} \qquad \forall\; i\neq j \tag{2.23}$$

$$\text{Cross}-\text{price elasticity of demand}: \quad \frac{\partial \log q_i}{\partial \log p_j} \; = \; -\frac{p_j}{q_i}\cdot(\mathbf{I}+\mathbf{\Sigma})^{-1}_{ij} \tag{2.24}$$

It is worth stopping to inspect equation (2.24) more closely. The first thing can notice is that the cross-price demand elasticities depend on the inverse $(\mathbf{I}+\mathbf{\Sigma})^{-1}$. This implies that, while cosine similarities are positive by construction, it is entirely possible for goods to be complements. This property of the model is discussed at length in Section 5.

Next, let us consider the case $i=j$, where (2.24) simply becomes the own residual demand elasticity. The first major difference between the GHL demand system and CES is that, while in CES the own demand elasticity is equal to a constant, here the own demand elasticity is an equilibrium object (as it depends on $\mathbf{q}$) and will generally differ among firm pairs. This implies that, unlike CES, this demand system produces heterogenous markups. In fact, we can see that two forces drive cross-sectional differences in market power across firms. The more familiar one is the incomplete passthrough from marginal cost to prices: that is, larger firms (high q_i) charge higher markups. The second force, which is instead a feature of hedonic demand models, is asymmetric product differentiation. That is, firms that produce "unique" products, as measured by the term $(\mathbf{I}+\mathbf{\Sigma})^{-1}_{ii}$, face a less elastic residual demand.

Next, I define the economic profits π_i as follows:

$$\begin{aligned}\pi_i(\mathbf{q}) \quad &\stackrel{\text{def}}{=} \quad p_i(\mathbf{q})\cdot q_i - h_i \\ &= \quad q_i(b_i-c_i)-\left(1+\frac{\delta_i}{2}\right)q_i^2-\sum_{j\neq i}\sigma_{ij}q_iq_j\end{aligned}$$

Firms compete à la Cournot: each firm i strategically chooses its output volume q_i by taking as given the output of all other firms. By taking the profit vector as a payoff function and the vector of quantities produced $\mathbf{q}$ as a strategy profile, I have implicitly defined a *linear-quadratic network game* (Ballester, Calvó-Armengol and Zenou, 2006, henceforth BCZ). The reason is that the matrix $\mathbf{\Sigma}$ can be conceptualized as the adjacency matrix of a weighted network: in this specific instance, it is the network of product market rivalry relationships that exists among the firms, based on the substitutability of their products.

Linear-quadratic network games belong to a larger class of games known as "potential games" (Monderer and Shapley, 1996): the key feature of potential games is that they can be described by a scalar function $\Phi(\mathbf{q})$, which we call the game's *potential*. The potential function can be thought of, intuitively, as the objective function of the *pseudo-planner* problem that is solved by the Nash equilibrium allocation. The potential function is shown below, together with the aggregate profit function $\Pi(\mathbf{q})$ and the aggregate welfare function $W(\mathbf{q})$:

$$\begin{aligned} \text{Aggregate Profit}: \quad \Pi(\mathbf{q}) &= \mathbf{q}'(\mathbf{b}-\mathbf{c}) \quad -\mathbf{q}'\left(\mathbf{I}+\frac{1}{2}\boldsymbol{\Delta}+\boldsymbol{\Sigma}\right)\mathbf{q} \\ \text{Cournot Potential}: \quad \Phi(\mathbf{q}) &= \mathbf{q}'(\mathbf{b}-\mathbf{c}) \quad -\mathbf{q}'\left(\mathbf{I}+\frac{1}{2}\boldsymbol{\Delta}+\frac{1}{2}\boldsymbol{\Sigma}\right)\mathbf{q} \\ \text{Total Surplus}: \quad W(\mathbf{q}) &= \mathbf{q}'(\mathbf{b}-\mathbf{c}) \quad -\frac{1}{2}\cdot\mathbf{q}'(\mathbf{I}+\boldsymbol{\Delta}+\boldsymbol{\Sigma})\mathbf{q} \end{aligned} \tag{2.25}$$

$$\text{where} \qquad \boldsymbol{\Delta} \overset{\text{def}}{=} \begin{bmatrix} \delta_1 & 0 & \cdots & 0 \\ 0 & \delta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_n \end{bmatrix} \tag{2.26}$$

The three functions in equation (2.25) are visually similar to each other; they only differ by the scalar weight applied to the quadratic terms. The Cournot potential Φ is somewhat of a hybrid between the aggregate profit Π and the total surplus W: the diagonal elements of the quadratic term are the same as the aggregate profit function, while the off-diagonal terms are the same as the aggregate surplus function. By maximizing the potential $\Phi(\mathbf{q})$, we find the Cournot-Nash equilibrium. I shall assume all these three functions are concave. Because the oligopolists in this model will be actual firms in the data (who produce positive output by definition) we can look directly at the unique internal solution.

Proposition 1. *The Cournot-Nash equilibrium of the game described above is* $\mathbf{q}^{\Phi}$ *– the maximizer of the potential function* $\Phi(\cdot)$*:*

$$\mathbf{q}^{\Phi} \quad \overset{\text{def}}{=} \quad \arg\max_{\mathbf{q}} \Phi(\mathbf{q}) \quad = \quad (2\mathbf{I}+\boldsymbol{\Delta}+\boldsymbol{\Sigma})^{-1}(\mathbf{b}-\mathbf{c}) \tag{2.27}$$

Proof. The derivation of the potential function, as well as the proof that its maximizer $\mathbf{q}^{\Phi}$ is the genuine Nash equilibrium, appear in Appendix A. □

Equation (2.27), which characterizes the Cournot-Nash equilibrium, tells us which factors determine the size of each firm in equilibrium. The diagonal matrix $\boldsymbol{\Delta}$, which contains the slopes of the marginal cost functions, captures economies of scale. $\boldsymbol{\Sigma}$ is the adjacency matrix of the network of product rivalries. $\mathbf{b}$ and $\mathbf{c}$ are, respectively, the demand and supply function intercepts. Hence, $(b_i - c_i)$ is simply the marginal surplus of the very first unit produced by firm i; also, b_i can be interpreted as a measure of vertical product differentiation (quality).

BCZ show that another way to interpret equation (2.27) is as a measure of *network centrality* – specifically, that developed by Katz (1953) and Bonacich (1987). The intuition is that firms that are more "isolated" in the network of product similarities face less product market competition and behave more like monopolists.

Centrality measures are a recurring feature of the literature on networks in macroeconomics (see Carvalho and Tahbaz-Salehi, 2019). In Appendix B, I discuss in further detail the link between Nash equilibrium and network centrality.

The discrepancy between the potential function and the total-surplus function implies that the network Cournot game delivers an equilibrium allocation that is not socially-optimal. A benevolent social planner can theoretically improve on the market outcome for two reasons. First, they can coordinate output choices across firms; second, they can internalize consumer surplus.

2.3. Separability of Consumer Surplus

Under GHL demand, consumer surplus has a desirable property, which I call *additive separability*: it means we can attribute a certain share of the consumer surplus to each firm. I will use this separability property to propose a measure of oligopoly power that varies by firm, and which I will be able to link to surplus appropriation.

Definition 3 (Additive Separability of Consumer Surplus)**.** Assume that the allocation $\mathbf{q}$ maximizes the consumer utility given the price vector $\mathbf{p}$. We say that the consumer surplus $S(\mathbf{q})$ is *additively separable* if it can be written as the sum over the set of firms of some function $s(b_i, q_i, p_i)$ that only depends on the triple (b_i, q_i, p_i). That is:

$$S(\mathbf{q}) = \sum_i s(b_i, q_i, p_i) \tag{2.28}$$

Proposition 2. *The consumer surplus function $S(\mathbf{q})$ from equation* (2.18) *is additively separable.*

Proof. Noting that the inverse demand function can be rearranged as $(\mathbf{I}+\boldsymbol{\Sigma})\mathbf{q} = \mathbf{b} - \mathbf{p}$, we can write equation (2.18) as:

$$\begin{aligned} S(\mathbf{q}) &= \mathbf{q}'(\mathbf{b}-\mathbf{p}) - \frac{1}{2}\mathbf{q}'(\mathbf{I}+\boldsymbol{\Sigma})\mathbf{q} \\ &= \mathbf{q}'(\mathbf{b}-\mathbf{p}) - \frac{1}{2}\mathbf{q}'(\mathbf{b}-\mathbf{p}) = \frac{1}{2}\mathbf{q}'(\mathbf{b}-\mathbf{p}) \end{aligned} \tag{2.29}$$

the last term can be rewritten in summation form as $\sum_i \frac{1}{2} q_i (b_i - p_i)$. □

By substituting p_i with the inverse demand function, we obtain the firm-level consumer surplus, which attributes to each firm i a certain share s_i of the consumer surplus $S(\mathbf{q})$:

$$s_i(\mathbf{q}) \stackrel{\text{def}}{=} \frac{1}{2}\left(q_i^2 + \sum_{j\neq i}\sigma_{ij} q_i q_j\right) \tag{2.30}$$

We can also define a firm-level *total* surplus function, which specifies for every firm i a certain share w_i of the total surplus $W(\mathbf{q})$:

$$\begin{aligned} w_i(\mathbf{q}) &\stackrel{\text{def}}{=} \pi_i(\mathbf{q}) + s_i(\mathbf{q}) \\ &= q_i(b_i - c_i) - \frac{1}{2}\left[(1+\delta_i) q_i^2 + \sum_{j\neq i}\sigma_{ij} q_i q_j\right] \end{aligned} \tag{2.31}$$

2.4. Oligopoly Power and Surplus Appropriation at the Firm-Level

The canonical oligopoly model with perfectly-substitutable products establishes the Herfindahl-Hirschmann Index (HHI) as a measure of market power. The reason is that the HHI relates the (market share-)weighted average of the inverse demand elasticities to the industry-wide inverse demand elasticity. Let $Q = \sum_i q_i$. Then:

$$\begin{aligned} \frac{\partial \log p}{\partial \log q_i} &= \frac{\partial \log p}{\partial \log Q} \cdot \frac{q_i}{Q} \\ \frac{q_i}{Q} \cdot \frac{\partial \log p}{\partial \log q_i} &= \frac{\partial \log p}{\partial \log Q} \cdot \left(\frac{q_i}{Q}\right)^2 \\ \sum_i \frac{q_i}{Q} \cdot \frac{\partial \log p}{\partial \log q_i} &= \frac{\partial \log p}{\partial \log Q} \cdot \mathrm{HHI} \end{aligned} \tag{2.32}$$

where $\mathrm{HHI} = \sum_i \left(\frac{q_i}{Q}\right)^2$ is the Herfindahl-Hirschmann Index. The lemma above illustrates that the reason why the HHI is informative about the residual demand elasticity is that the individual *market shares* are informative about demand elasticity (this fact is frequently forgotten). The first line of equation (2.32) evinces this: the individual inverse demand elasticity is simply equal to the industry-level inverse demand elasticity, times the market share of firm i. Hence, if we wanted to derive a firm-level counterpart of the HHI index, it would simply be the market share of firm i.

Let us now return to the network Cournot model and define the following statistic for firm i.

Definition 4. I define ω_i, the *weighted market share* of firm i, as follows:

$$\omega_i \stackrel{\text{def}}{=} \frac{q_i}{q_i + \sum_j \sigma_{ij} q_j} \tag{2.33}$$

Notice that, under homogenous products ($\sigma_{ij} = 1 \; \forall \, i, j$) this is simply the market share of firm i . It is possible to show that the ratio of firm profits π_i to consumer surplus s_i is proportional to the weighted market share ω_i.

Lemma 1. *In the Cournot-Nash equilibrium allocation, the ratio of profits to consumer surplus for firm i is proportional to its weighted market share - specifically:*

$$\frac{\pi_i}{s_i} = (2 + \delta_i)\, \omega_i \tag{2.34}$$

Proof. See Appendix K. □

Therefore, in the Network Cournot model, the similarity-weighted market share ω_i replaces the HHI as a firm-level measure of market power that accounts for product differentiation. As is the case for the HHI, the weighted market share is an equilibrium object—an endogenous outcome of the Cournot game played by the oligopolists.

The identity in Lemma (1) reflects the fact that, in my model, there are no clearly-defined industry boundaries. This is also the case in the real world: if we consider antitrust lawsuits for example, a major object of litigation is the market's definition. Defendants (alleged monopolies) have an incentive to define the relevant market broadly, while plaintiffs have an incentive to define the relevant market narrowly.

In my model, firms exist in a continuous space of product characteristics. Hence, there is no uniquely-defined peer group that we can compare each firm to. To understand how dominant firm i is, we need to compare its market share vis-à-vis every other firm in the economy, weighting each of them by their distance in the space of product characteristics.

The Herfindahl Index can be seen as a special case of the weighted market share: as a measure of surplus appropriation, it is only valid in the special case where similarity scores are dichotomous (implying sharp industry boundaries) and firms are exchangeable ($b_i - c_i$ is constant across firms).

2.5. Market Structure Counterfactuals

A key application of my theoretical model is to study how welfare statistics - such as total surplus - respond to changes in market structure. What that means is that, having made the required assumption that firms compete by a well-defined set of rules (thus far we have assumed Cournot oligopoly), we can then consider counterfactuals in which the same firms play by a different set of rules. In this Subsection, I define four of these counterfactuals: each of these counterfactuals corresponds to the solution of a specific maximization problem.[5]

The first counterfactual that I consider is *perfect competition*: firms act as atomistic producers, and price all units sold at marginal cost.

Definition 5. The *Perfect Competition* allocation $\mathbf{q}^W$ is defined as the maximizer of the aggregate total surplus function $W(\mathbf{q})$:

$$\mathbf{q}^W \quad \stackrel{\text{def}}{=} \quad \arg\max_{\mathbf{q}} W(\mathbf{q}) \quad = \quad (\mathbf{I} + \mathbf{\Delta} + \mathbf{\Sigma})^{-1} (\mathbf{b} - \mathbf{c}) \tag{2.35}$$

The second counterfactual that I consider is called *Monopoly*: it represents a situation in which one agent (that does not internalize consumer surplus) has control over all the firms in the economy and maximizes aggregate profits.

Definition 6. The *Monopoly* allocation is defined as the maximizer of the aggregate profit function $\Pi(\mathbf{q})$:

$$\mathbf{q}^\Pi \quad \stackrel{\text{def}}{=} \quad \arg\max_{\mathbf{q}} \Pi(\mathbf{q}) \quad = \quad (2\mathbf{I} + \mathbf{\Delta} + 2\mathbf{\Sigma})^{-1} (\mathbf{b} - \mathbf{c}) \tag{2.36}$$

This allocation can be alternatively conceptualized as an economy with no antitrust policy, where firms have unlimited ability to coordinate their supply choices.

While the *Monopoly* counterfactual is an interesting limit case, using the model we can also study the welfare impact of mergers and collusion among specific firms.

When it comes to modeling mergers and collusions, the I.O. literature has used multiple approaches. Following Baker and Bresnahan (1985), I choose model mergers and collusion interchangeably as *coordinated pricing*. That is, I assume that the merger or the collusion does not affect the product range offered by

[5] The closed-form expressions for the output vector $\mathbf{q}$ which I provide below assume an internal solution. For my empirical analysis, I also compute a numerical solution that is subject to a non-negativity constraint on $\mathbf{q}$ and I verify it is approximately equal to the unconstrained solution (error $< 0.1\%$ for the total surplus function in Perfect Competition). The non-negativity constraint binds for very few firms.

the merging/colluding enterprises; instead, a single agent determines the output of the merging firms to maximize the joint profits.[6]

Lemma 2. *Consider, without loss of generality, a merger or collusion between companies* $\{1, 2, ..., m\}$*; then, partition the matrix* $\mathbf{\Sigma}$ *by separating the first* m *rows and columns as follows:*

$$\mathbf{\Sigma} \;=\; \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix} \tag{2.37}$$

The post-merger equilibrium allocation maximizes the following modified potential function:

$$\Phi\left(\mathbf{q}\right) \;=\; \mathbf{q}'\left(\mathbf{b}-\mathbf{c}\right) - \mathbf{q}'\left(\mathbf{I}+\frac{1}{2}\mathbf{\Delta}\right)\mathbf{q} - \frac{1}{2}\cdot\mathbf{q}'\begin{bmatrix} 2\mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}\mathbf{q} \tag{2.38}$$

Proof. See Appendix K. □

The maximizer of the re-defined $\Phi\left(\mathbf{q}\right)$, which corresponds to the post-merger equilibrium allocation, is:

$$\mathbf{q}^{\Phi} = \left(2\mathbf{I}+\mathbf{\Delta}+\begin{bmatrix} 2\mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}\right)^{-1}\left(\mathbf{b}-\mathbf{c}\right) \tag{2.39}$$

That is, to simulate the new equilibrium following a merger or a collusion among existing firms, one only needs to amend the potential function by doubling the off-diagonal quadratic terms corresponding to the merging firms. It is easily verified that when all firms are merged, $\Phi\left(\mathbf{q}\right)$ simply becomes the aggregate profit function $\Pi\left(\mathbf{q}\right)$, and the equilibrium allocation converges to the Monopoly counterfactual (equation 2.36).

Another interesting counterfactual is one in which resources are allocated efficiently but the labor supply is fixed. That is, the social planner maximizes the aggregate surplus function subject to the constraint of using no more labor than in the observed Cournot equilibrium.

Definition 7. I define the resource-efficient counterfactual $\mathbf{q}^H$ as the solution to the following constrained maximization problem:

$$\mathbf{q}^{H} \quad \stackrel{\text{def}}{=} \quad \arg\max_{\mathbf{q}} W\left(\mathbf{q}\right) \quad \text{s.t.} \quad H\left(\mathbf{q}\right) = H\left(\mathbf{q}^{\Phi}\right) \tag{2.40}$$

Setting up the Lagrangian and using $(1-\mu)$ as the Lagrange multiplier, we find that the resource-efficient counterfactual takes the form:

$$\mathbf{q}^{H} \;=\; \left(\mathbf{I}+\mu\mathbf{\Delta}+\mathbf{\Sigma}\right)^{-1}\left(\mathbf{b}-\mu\mathbf{c}\right) \tag{2.41}$$

where μ solves:

$$H\left(\mathbf{q}^{H}\left(\mu\right)\right) \;=\; H\left(\mathbf{q}^{\Phi}\right) \tag{2.42}$$

The Lagrange multiplier term μ turns out to be the common markup charged by all firms in the resource-efficient counterfactual.

Lemma 3. *The Resource-efficient counterfactual* $\mathbf{q}^H$ *equalizes markups across firms.*

[6]This approach is particularly attractive, from a tractability standpoint, in a setting like this, where products are differentiated. The alternative would be to make some heroic assumptions about the nature and characteristics of the product supplied by the combined entity.

Proof. Let all firms price at a constant markup μ over marginal cost:

$$p_i \;=\; \mu \cdot \mathrm{MC}_i \tag{2.43}$$

expanding the expression for the marginal cost and the equilibrium price we have:

$$\mathbf{b} - (\mathbf{I} + \mathbf{\Sigma})\,\mathbf{q} \;=\; \mu\,(\mathbf{c} + \mathbf{\Delta q}) \tag{2.44}$$

rearranging the equation above we obtain (2.41). □

Because this counterfactual uses the same amount of labor as the observed equilibrium, by comparing welfare in this allocation to the first-best we can effectively break down the deadweight loss into two components – one linked to misallocation, the other linked to labor suppression. We can also interpret this counterfactual as the deadweight loss in an alternative model where the supply of labor is completely inelastic. Notice that when this allocation is not constrained by the labor supply (the Lagrange multiplier $1-\mu$ is zero), the common markup is one (firms price at marginal cost) and the resource-efficient allocation coincides with perfect competition.

The counterfactuals considered thus far do not account for how a firm's incentives to participate in the market are affected by the intensity of competition. When the market moves from Cournot competition to (say) Bertrand [7] or perfect competition, the resulting lower profits might insufficient to cover fixed costs, and therefore too low to justify a firm's continued existence. If this is the case, perfect competition may not be a realistic benchmark in the long-run: this is the classical criticism of static welfare analysis.

Next, I construct an "efficient" allocation that takes into account (to the extent possible in a static model) these dynamic incentives. The starting point is again a benevolent social planner, to which we are adding a constraint, in the form of a participation condition on the firms' side: firms have to be able (on average) to recover their fixed costs (F) at the optimum.[8]

Definition 8. The Second-Best Allocation $\mathbf{q}^{2nd}$ is defined as the solution to the following constrained maximization problem:

$$\mathbf{q}^{2nd} \quad \stackrel{\text{def}}{=} \quad \arg\max_{\mathbf{q}} W(\mathbf{q}) \quad \text{s.t.} \quad \Pi(\mathbf{q}) \geq F \tag{2.45}$$

where

$$F \stackrel{\text{def}}{=} \sum_i f_i \tag{2.46}$$

Setting up the Lagrangian of this problem and imposing λ as the Lagrangian multiplier, we find that the resource-efficient counterfactual takes the form:

$$\mathbf{q}^{2nd} \quad = \quad \left[\frac{1+2\lambda}{1+\lambda}\cdot(\mathbf{I}+\mathbf{\Sigma})+\mathbf{\Delta}\right]^{-1}(\mathbf{b}-\mathbf{c}) \tag{2.47}$$

[7] The Bertrand model is covered in Appendix F.

[8] There are two reasons why I consider a constraint on aggregate profits rather than individual profits ($\pi_i \geq f_i$). The first is that such individual constraint is already violated by many firms in the observed (Cournot) equilibrium. The second reason is that adding individual constraints for each firms would make the optimization problem numerically intractable, since we would need to solve for thousands of constraints (one for each firm in the model).

Assuming that the constraint binds at the optimum, the Lagrange multiplier λ solves:

$$\Pi\left(\mathbf{q}^{2nd}\left(\lambda\right)\right) \;=\; F \tag{2.48}$$

As the constraint is relaxed ($\lambda \to 0$), this counterfactual allocation converges to the first-best. When the constraint becomes arbitrarily tight ($\lambda \to \infty$), it converges to the *Monopoly* allocation.

In addition to the counterfactuals considered above, which admit closed-form solutions, we can simulate the introduction or the removal of granular firms. The latter can be trivially implemented by computing an allocation where a firm's output is constrained to be zero. In order to simulate instead the introduction of new firms, we require additional assumptions or data. Namely, in order to simulate the introduction of an additional firm (let us label it firm zero), we would need to know the value of $(b_0 - c_0)$, as well as the firm's similarity to every other firm in the economy ($\mathbf{a}_{i0}$). One such counterfactual is considered in Section (4).

2.6. Adding a Continuum of firms with Endogenous Entry

Next, I show how to expand the model to include a continuum of atomistic firms that behave competitively and can enter and exit endogenously. This extension of the model allows me to accomplish two things: 1) incorporate firms for whom we do not observe product similarity data – that is, foreign and private firms; 2) it allows to incorporate entry and exit in an otherwise static model. The idea is that we can model unobserved companies as atomistic firms.

The key to tractably integrating these atomistic firms in the model is an aggregation result. I describe these atomistic firms through a productivity distribution: the set of active atomistic firms is then characterized by a productivity cut-off value, in the style of Hopenhayn (1992).

Next, I show that these atomistic companies can be aggregated into a representative firm: variations in the size of the representative firm reflect the intensive margin of production as well as the extensive margin (the entry/exit of the atomistic firms). I index this representative firm $i = n+1$, effectively adding a row and a column to the matrices $\mathbf{A}'\mathbf{A}$ and $\mathbf{\Delta}$ and adding one dimension to the vector $\mathbf{b}$.

Proposition 3. *Assume that there is a mass one of potential entrants that are indexed by a productivity parameter $z \in (\underline{z}, \infty)$ and that produce a homogeneous good using the following quadratic cost function:*

$$h\left(z\right) \;=\; \frac{\delta\left(z\right)}{2} \cdot q^2\left(z\right) \tag{2.49}$$

with $\underline{z} > 0$ and

$$\delta\left(z\right) \;=\; \frac{1}{z} \tag{2.50}$$

Assume also that the firms face cost of entry equal to one unit of labor and that the probability density of type-z potential entrants is given by

$$f\left(z\right) \;=\; \frac{\beta - 1}{z^{\beta+1}} \tag{2.51}$$

implying that z follows a Pareto distribution with shape parameter β and scale parameter $\underline{z} \stackrel{\text{def}}{=} [(\beta-1)/\beta]^{\frac{1}{\beta}}$.[9] *Then, as the parameter β converges down to 1, the cost function of the corresponding aggregate representative firm is approximated by*

$$h_{n+1} \;=\; \frac{q_{n+1}^2}{2} \tag{2.52}$$

where and h_{n+1} and q_{n+1} are, respectively, the labor input and the output of the representative firm, and the productivity cutoff for entry converges to $z_{\min} = \frac{1}{q_{n+1}}$.

Proof. See Appendix K. □

Because employment and revenues are proportional to z, it follows that, if the assumptions above are respected, both the revenue and employment distribution of firms also approximate a Pareto distribution with shape parameter $\beta = 1$, sometimes called a Zipf Law.

Although this might look like a knife-edge assumption, it is not. It is a well-documented empirical regularity that the size distribution of firms closely approximates a Pareto distribution with shape parameter $\beta = 1$. This stylized fact was confirmed to hold for both the employment *and* the revenue distribution of US firms by Axtell (2001), using Census micro-data.

Because the representative firm behaves competitively, its first order condition will differ from that of granular firms $\{1, 2, ..., n\}$. The latter maximize individual profits:

$$\pi_i'(q_i) \;=\; 0 \qquad \text{for} \quad i = 1, 2, ..., n \tag{2.53}$$

The representative firm, on the other hand, prices at marginal cost, and therefore maximizes total surplus:

$$W'(q_i) \;=\; 0 \qquad \text{for} \quad i = n+1 \tag{2.54}$$

We can write the full system of first order conditions in linear algebra notation as:

$$0 \;=\; \begin{bmatrix} \mathbf{b}^{(n)} - \mathbf{c}^{(n)} \\ b_{n+1} - c_{n+1} \end{bmatrix} + \left(\begin{bmatrix} 2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} + \boldsymbol{\Sigma} + \boldsymbol{\Delta} \right) \begin{bmatrix} \mathbf{q}^{(n)} \\ q_{n+1} \end{bmatrix} \tag{2.55}$$

where $c_{n+1} = 0$, $\delta_{n+1} = 1$ and the superscript (n) identifies the sub-vector corresponding to the granular firms. A simpler way to rewrite this set of equations is

$$0 \;=\; \mathbf{b} - \mathbf{c} - (\mathbf{I} + \mathbf{G} + \boldsymbol{\Sigma} + \boldsymbol{\Delta})\,\mathbf{q} \tag{2.56}$$

where $\mathbf{G}$ is a diagonal matrix that identifies granular firms – that is, whose diagonal elements equal 1 for

[9] While the revenue and employment distribution of US firms approximates a Pareto Distribution with scale parameter equal to one (a Zipf Law), this distribution has the undesirable property that its mean (and therefore q_{n+1} and h_{n+1}) grows unboundedly as $\beta \to 1^+$. This particular choice of the scale parameter ensures that q_{n+1} and h_{n+1} integrate to a finite number as $\beta \to 1^+$.

firms 1 to n and to 0 for firm $n+1$:

$$\mathbf{G} \;=\; \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{2.57}$$

The potential function for the model that includes the representative firm is:

$$\Phi\left(\mathbf{q}\right) \;=\; \mathbf{q}'\left(\mathbf{b}-\mathbf{c}\right)-\frac{1}{2}\mathbf{q}'\left(\mathbf{I}+\mathbf{G}+\mathbf{\Sigma}+\mathbf{\Delta}\right)\mathbf{q} \tag{2.58}$$

and the equilibrium quantity vector is:

$$\mathbf{q}^{\Phi} \;=\; \left(\mathbf{I}+\mathbf{G}+\mathbf{\Delta}+\mathbf{\Sigma}\right)^{-1}\left(\mathbf{b}-\mathbf{c}\right) \tag{2.59}$$

3. Data and Identification/Calibration

In this section, I outline the data used to estimate the model in Section 2. Additional details are provided in Appendix D, which also contains Table 2, where model mapping and identification are summarized.

3.1. Firm Financials

My data source for firm financials is the Compustat database, which I access via the Wharton Research Data Services (WRDS) platform. From this database, I extract information on firm revenues, Costs of Goods Sold (COGS), Selling General and Administrative (SGA) costs, R&D expenditures and Property Plant and Equipment (PPE).

I follow (De Loecker, Eeckhout and Unger, 2020, henceforth DEU) in mapping accounting revenues to model revenues, COGS to variable costs, and in computing an estimate of fixed costs costs (f_i):

$$f_i \;\leftarrow\; \text{SGA}_i+\text{Property Plant \& Equipment}_i\times\text{User Cost of Capital} \tag{3.1}$$

3.2. Text-Based Product Similarity

The key data ingredient that we need, in order to estimate my model, is the matrix of product similarities $\mathbf{A}'\mathbf{A}$. The empirical counterpart of this object is provided by Hoberg and Phillips (2016, henceforth HP).

HP created a publicly-available database that provides product cosine similarities for the universe of public corporations in the United States. These cosine similarities originate from natural language processing (NLP) of 10-K filings, and are time-varying. A complete matrix of similarities is provided for every year, beginning in 1997.

The 10-K is a mandatory form that is filed by American public corporations with the U.S. Securities and Exchange Commission on a yearly basis. Item 1 of the 10-K is a long and detailed description of the product

or service sold by the company. HP's product cosine similarities are constructed by comparing these textual product descriptions.

I briefly outline the construction of this dataset. HP start by building a vocabulary of 61,146 words that firms use to describe the characteristics of their products.[10] Based on this vocabulary, HP produce, for each firm i, a vector of word frequencies $\mathbf{o}_i$. Each of component of this vector corresponds to a word in HP's vocabulary, and is equal to the number of times that word appears in firm i's 10-K product description:

$$\mathbf{o}_i = \begin{bmatrix} o_{i,1} \\ o_{i,2} \\ \vdots \\ o_{i,61146} \end{bmatrix} \tag{3.2}$$

Similar to the model in Section 2, this vector is then normalized (divided by the Euclidean norm). We have thus obtained the empirical counterpart of $\mathbf{a}_i$:

$$\mathbf{a}_i = \frac{\mathbf{o}_i}{\|\mathbf{o}_i\|} \tag{3.3}$$

finally, all $\mathbf{a}_i$ vectors are dot-multiplied to obtain $\mathbf{A}'\mathbf{A}$:

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} \mathbf{a}_1'\mathbf{a}_1 & \mathbf{a}_1'\mathbf{a}_2 & \cdots & \mathbf{a}_1'\mathbf{a}_n \\ \mathbf{a}_2'\mathbf{a}_1 & \mathbf{a}_2'\mathbf{a}_2 & \cdots & \mathbf{a}_2'\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n'\mathbf{a}_1 & \mathbf{a}_n'\mathbf{a}_2 & \cdots & \mathbf{a}_n'\mathbf{a}_n \end{bmatrix} \tag{3.4}$$

Hence, to the extent that word frequencies are a good proxy for characteristics, the resulting matrix is the exact empirical counterpart to $\mathbf{A}'\mathbf{A}$ in my theoretical model. The fact that all publicly-traded firms in the United States are required to file a 10-K form makes the this data set unique in that it covers the near entirety (97.8%) of the Compustat universe.

HP use these cosine similarities to produce a dynamic industry classification, called TNIC, which they extensively validate: one way they validate their data (in the paper that presents their methodology) is by using another dataset called CapitalIQ. This dataset provides dummy variables for a sub-set of Compustat firm pairs which identify product market rivalry relationships; they are based on corporate filings as well as other sources (no time variation is available in this dataset). HP show that TNIC outperforms SIC and NAICS in predicting competitor pairs in CapitalIQ.

Since their introduction in 2011, HP's industry classifications have become standard in the empirical corporate finance literature, where they have replaced NAICS and SIC for a variety of applications. A major reason for this methodological shift is that HP's dataset addressed an important limitation of traditional industry classifications. While these have often been used (for lack of better alternatives) to capture product

[10] I report here verbatim the methodology description from the original paper by Hoberg and Phillips (2016):"[...] *In our main specification, we limit attention to nouns (defined by Webster.com) and proper nouns that appear in no more than 25 percent of all product descriptions in order to avoid common words. We define proper nouns as words that appear with the first letter capitalized at least 90 percent of the time in our sample of 10-Ks. We also omit common words that are used by more than 25 percent of all firms, and we omit geographical words including country and state names, as well as the names of the top 50 cities in the United States and in the world.* [...]"

Figure 2: Network Visualization of the Hoberg-Phillips Dataset

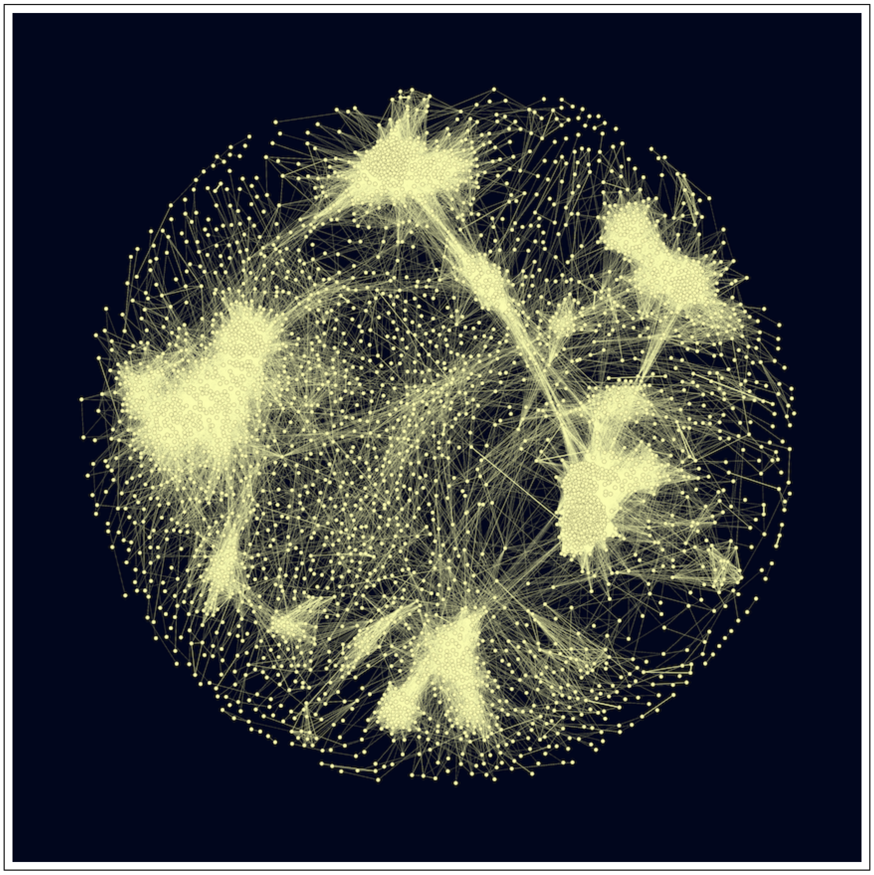


Figure Notes: The following diagram is a two-dimensional representation of the network of product similarities computed by Hoberg and Phillips (2016), which is used in the estimation of the model presented in Section 2. The data covers the universe of Compustat firms in 2004. Firm pairs that have thicker links are closer in the product market space. These distances are computed in a space that has approximately 61,000 dimensions. To plot this high-dimensional object over a plane, I applied the gravity algorithm of Fruchterman and Reingold (1991), which is standard in social network analysis.

market competition[11], it is well-known that they are based on the concept of *production process* similarity, not product similarity[12]. This is also one reason why, in the I.O. and Antitrust literature, NAICS and SIC are generally only used to estimate production functions[13].

There are other factors that differentiate HP's database from traditional industry classifications. While NAICS and SIC are binary (firms are either in the same industry or different industries), HP's database also provides continuous similarity scores ranging from zero to one, thus accommodating the inherent fuzziness of product market rivalries. While NAICS and SIC are seldom updated, HP's similarity scores are updated

[11]Before HP's data was published, Bloom, Schankerman and Van Reenen (2013) constructed cosine similarities to estimate R&D spillovers. They used Compustat Segments data, which is based on NAICS/SIC industries. This dataset's coverage of Compustat is insufficient to estimate my model (not enough firms/years available).

[12]See the following Bureau of Labor Statistics Guide.

[13]For example: DEU's method to compute markups uses production function estimates for NAICS industries.

yearly. While NAICS and SIC are arbitrarily assigned (Chen et al., 2016 show that firms strategically manipulate their industry classifications), HP's similarity scores are rule-driven and incentive-compatible: executives face legal liability for misrepresenting company information in SEC filings.

I begin my empirical analysis by visualizing HP's dataset. To do so, I have to reduce the dimensionality of the dataset from 61,146 (the number of words in the HP's vocabulary) to two. I do so using the algorithm of Fruchterman and Reingold (1991, henceforth FR), which is widely used in network science to visualize weighted networks[14].

The result of this exercise is Figure 2: every dot in the graph is a publicly traded firm as of 2004. Firm pairs that have a high cosine similarity appear closer, and are joined by a thicker line. Conversely, firms that are more dissimilar are not joined, and are more distant. From the graph, we can see that the distribution of firms over the space of product characteristics is manifestly uneven: some areas are significantly more densely populated with firms than others. Also, the network displays a pronounced community structure: large groups of firms tend to cluster in certain areas of the network.

In Appendix C, I show that this visualization is not an artifact of dimensionality reduction or measurement error: notwithstanding the dimensionality reduction, a remarkable degree of overlap exists between the macro-clusters of this network and broad economic sectors. In addition, this exercise allows me to independently validate HP's product similarity data.

3.3. Identification of Output, Prices and Cost Intercept

All of the unobserved variables in the model are identified subject to two parameters: (α) which controls the degree of horizontal differentiation between goods; and the diagonal matrix ($\mathbf{\Delta}$), which controls returns to scale. I will first show how to identify the remaining variables conditional on these two parameters and then I will illustrate my calibration procedure for α and $\mathbf{\Delta}$.

I measure the matrix $\mathbf{A}'\mathbf{A}$ using Hoberg and Phillips (2016)'s cosine similarity data. Given α, this provides an estimate of $\mathbf{\Sigma}$. We can then identify real output q_i from revenues and total variable cost data:

$$q_i \;=\; \sqrt{\frac{\pi_i}{1+\delta_i/2}} \qquad \text{if} \quad i \le n \tag{3.5}$$

If the model includes a representative competitive firm $n+1$, the identification of q_i for this firm will be different. Specifically, the marginal cost pricing condition ($p_{n+1} = \mathrm{MC}_{n+1}$) implies that:

$$q_i \;=\; \sqrt{\pi_i + h_i} \qquad \text{if} \quad i = n+1 \tag{3.6}$$

where $(\pi_{n+1} + h_{n+1})$ is measured as the Gross Value Added of private and foreign firms, which I compute using the OECD Trade in Value Added (TiVA) Dataset.

[14]The algorithm models the network nodes as particles, letting them dynamically arrange themselves on a bidimensional surface as if they were subject to attractive and repulsive forces. One known shortcoming of this algorithm is that it is sensitive to the initial configurations of the nodes, and it can have a hard time uncovering the cluster structure of large networks. To mitigate this problem, and to make sure that the cluster structure of the network is properly displayed, I pre-arrange the nodes using the OpenOrd algorithm (which was developed for this purpose) before running FR.

Having identified q_i, we can then pin down the vector of prices and the cost function intercepts:

$$p_i = \frac{p_i q_i}{q_i} \qquad c_i = \frac{h_i}{q_i} - \frac{\delta_i}{2} q_i \tag{3.7}$$

Finally, I identify the demand intercept b_i using equation (2.27):

$$\mathbf{b} = (2\mathbf{I} + \boldsymbol{\Delta} + \boldsymbol{\Sigma})\, \mathbf{q} + \mathbf{c} \tag{3.8}$$

or, in the presence of an representative competitive firm:

$$\mathbf{b} = (\mathbf{I} + \mathbf{G} + \boldsymbol{\Delta} + \boldsymbol{\Sigma})\, \mathbf{q} + \mathbf{c} \tag{3.9}$$

3.4. Calibration of α and $\boldsymbol{\Delta}$

The last step required to take the model to the data is to calibrate the scalar α and the diagonal matrix $\boldsymbol{\Delta}$. Let us start from the latter. To calibrate each diagonal element δ_i, we use the fact that the markup (price-marginal cost ratio) of firm i can be written as a function of observables (revenues, total variable costs) and δ_i – i.e. the markup μ_i is identified conditional δ_i.

Lemma 4. *The markup μ_i is equal to:*

$$\mu_i \stackrel{\text{def}}{=} \frac{p_i}{\text{MC}_i} = \frac{(2 + \delta_i) \cdot p_i q_i}{2 \cdot h_i + \delta_i \cdot p_i q_i} \tag{3.10}$$

Proof. See Appendix (K). □

DEU compute the revenue-weighted average markup for the same universe of companies (Compustat). My strategy for calibrating $\boldsymbol{\Delta}$ is to target their estimate. The detailed methodology for calibrating $\boldsymbol{\Delta}$ is outlined in detail in Appendix D.

To calibrate α, we rewrite equation (2.21) as:

$$\left| \frac{\partial \log p_i}{\partial \log q_j} \right| = \alpha \cdot \mathbf{a}_i' \mathbf{a}_j \frac{q_j}{p_i} \quad \forall\, i \neq j \tag{3.11}$$

By calibrating $\boldsymbol{\Delta}$, we have already pinned q_i and p_i. The matrix of cosine similarities of Hoberg and Phillips (2016) provides the empirical counterpart to $\mathbf{A}'\mathbf{A}$. Hence, the matrix of equilibrium cross-price demand elasticities is identified given α.

My strategy for calibrating α is to target target microeconometric estimates from the Industrial Organization literature. I obtain, for a number of firm pairs, estimates of the cross-price demand elasticity from empirical IO studies that estimate the demand function econometrically. These estimates of the cross-price demand elasticity are then manually matched to to the corresponding firm pair in Compustat. Finally, for each firm pair, I can obtain an estimate of α by rearranging equation (3.11):

$$\hat{\alpha}_{ij} = \left| \frac{\partial \log p_i}{\partial \log q_j} \right| \Big/ \left(\mathbf{a}_i' \mathbf{a}_j \frac{q_i}{p_i} \right) \tag{3.12}$$

In the absence of mis-specification and measurement error, all these estimates $\hat{\alpha}_{ij}$ would return the same

estimate. What I obtain in this case, instead, is a range of estimates. I calibrate α to the median value among these estimates, which is 0.05. The full methodology is presented in Appendix D, where I also discuss how the model fits non-targeted moments in the data.

4. Empirical Findings

In this section, I present the results of the estimation of my model. My baseline estimates reflect the model implementation that only includes granular firms (Compustat). In the next section, I discuss the robustness of my estimates to the inclusion of private and foreign firms as a continuum of atomistic firms that enter endogenously.

4.1. Welfare Statics

My first empirical exercise is to compute total surplus and to break it down into profits and consumer surplus. This is done for both the observed equilibrium (which is assumed to be a Nash-Cournot equilibrium) and the counterfactuals considered in Section 2. These estimates are all shown in Table 1.

I estimate that the publicly-traded firms earn an aggregate economic profit of \$5 trillion and produce an estimated total surplus of \$9.1 trillion. Consumer surplus is therefore estimated to be about \$4 trillion. About 55% of the total surplus produced is appropriated by the companies in the form of oligopoly profits. For context, the GDP of U.S. corporations in the same year (2017) is \$11 trillion[15].

The first counterfactual I consider, *Perfect Competition*, appears in the second column. By comparing this counterfactual with the Cournot-Nash allocation we can see that the welfare costs of oligopoly are significant. Under perfect competition, aggregate surplus is significantly higher – \$10.2 trillion – hence, the deadweight loss amounts to about 11% of the total surplus.

While the effects of oligopoly on Pareto efficiency are significant, perhaps even more significant are the *distributional* effects. When firms price at marginal cost, a much larger share of the surplus goes to the consumer: \$8.2 trillion, more than double than in the Cournot allocation. This amounts to 80% of the total surplus.

The next counterfactual I analyze, the *Monopoly* counterfactual, appears in the third column: it represents a scenario in which all firms are controlled by a single decision-maker that coordinates supply choices. In this allocation, aggregate surplus is significantly lower than in the Network Cournot equilibrium allocation: \$8.2 trillion. Despite the decrease in aggregate welfare, profits are markedly higher: \$5.7 trillion. Consequently, consumer surplus is reduced to just \$2.5 trillion, a mere 33% of the total. One interpretation of this exercise is that policies that prevent coordination between firms (antitrust) have a large positive impact on welfare.

Next, I consider the *Resource Efficient* counterfactual, in which the social planner maximizes total surplus subject to the constraint of not using more labor than the Cournot equilibrium.

[15] The difference between GDP and total surplus is that total surplus does not include the value of labor input but it does include the value of inframarginal consumption. GDP, on the other hand, includes the value of labor input but not the inframarginal value of consumption. In this model each unit of labor is paid exactly its marginal disutility, hence there is no inframarginal value of leisure.

Table 1: Welfare Statics (2017)

	Scenario	*Cournot-Nash* (1)	*Perfect Competition* (2)	*Monopoly* (3)	*Resource-Efficient* (4)	*Second-Best* (5)
Welfare Statistic	**Variable**	$\mathbf{q}^{\Phi}$	$\mathbf{q}^{W}$	$\mathbf{q}^{\Pi}$	$\mathbf{q}^{H}$	$\mathbf{q}^{2nd}$
Total Surplus (US$ trillions)	$W(\mathbf{q})$	9.086	10.208	8.183	9.869	10.017
Aggregate Profits (US$ trillions)	$\Pi(\mathbf{q})$	5.043	1.995	5.673	3.135	3.910
Consumer Surplus (US$ trillions)	$S(\mathbf{q})$	4.043	8.213	2.510	6.735	6.106
Total Surplus / Perfect Competition	$\frac{W(\mathbf{q})}{W(\mathbf{q}^{W})}$	0.890	1.000	0.802	0.967	0.981
Aggregate Profit / Total Surplus	$\frac{\Pi(\mathbf{q})}{W(\mathbf{q})}$	0.555	0.195	0.693	0.318	0.390
Consumer Surplus / Total Surplus	$\frac{S(\mathbf{q})}{W(\mathbf{q})}$	0.445	0.805	0.307	0.682	0.610

Table Notes: The following table shows my estimates of aggregate profits, consumer surplus and total surplus in each of the counterfactuals scenarios presented in Section 4.

In this scenario, markups across firms have been equalized, but not eliminated. By removing all dispersion in markups, this counterfactual targets the malallocative effects of oligopoly.

The total surplus produced in this counterfactual is \$9.9 trillion, about 3.3 percentage points lower than in perfect competition, and \$800 billion higher than the observed Cournot-Nash equilibrium. Most of the surplus produced – \$6.9 trillion, or 68% of the total – goes to the consumer; profits are reduced to \$3.1 trillion. Because labor is fixed, all the welfare gains with respect the Cournot equilibrium come from the *reallocation* of labor. Hence, an important take-away from this counterfactual is that a large share of the inefficiencies from oligopoly are driven by resource misallocation. A different way to say this is that the dispersion in markups (caused by oligopolistic competition with differentiated products) matters at least as much as the level of markups in determining the overall deadweight loss.

The last counterfactual that I consider is the *Second-Best,* in which a benevolent social planner maximizes aggregate surplus subject to an aggregate participation constraint (profits must cover fixed costs on average). In this counterfactual, total surplus is very close to the level achieved by the perfectly-competitive outcome: \$10 trillion. The main difference is the surplus split: the consumer receives \$6.1 trillion (two trillion less than under perfect competition but two more than under Cournot), while total profits amount to \$3.9 trillion.

In sum, my measurements suggest that the oligopoly power of U.S. public firms has significant consequences for aggregate welfare, and that it impacts consumer welfare through two channels: it increases the dispersion of markups, generating resource misallocation which raises the deadweight loss; it also increases the level of markups, which in turn affects how surplus is shared between producers and consumers.

4.2. Time Trends in Total Surplus and Consumer Surplus

HP's cosine similarity data is available starting from 1997. By mapping my model to Compustat data year by year, I can produce annual estimates of the welfare metrics previously presented. This allows me to study the welfare implications of the rising concentration of US industries. Most importantly, because my model leverages HP's time-varying product similarity data, these estimates account for how the product offering of US public firms changed over time. This is another contribution of this study.

In Figure 3, I plot aggregate consumer surplus S (the dark area) and profits Π (the light area) for every year between 1997 and 2017. The combined area represents total surplus W. I also plot, on the right axis (dotted black line), profits as a share of total surplus Π/W. All these statistics refer to the (observed) Cournot equilibrium.

The graph shows that the total surplus produced by US public corporations has nearly doubled between 1997 and 2017 from \$4.6 trillion to \$9.1 trillion between 1997 and 2017. Profits have increased more-than-proportionally with respect to consumer surplus – from about \$2.2 trillion to about \$5 trillion. Consumer surplus increased instead from \$2.2 trillion in 1997 to about \$4 trillion i 2017. As a consequence, the profit share of surplus has increased from about 50% of total surplus to nearly 56%. The consumer appears to capture a decreasing share of the surplus generated by public companies.

In Figure 4, I plot, over the same period, the percentage gain in total surplus from moving from the competitive equilibrium $\mathbf{q}^{\Phi}$ to the first best $\mathbf{q}^{W}$. This is the deadweight loss from oligopolistic behavior, and is plotted as the darker line. Its trend that mimic that of profit share of surplus: it increased from 8.5% (in 1997) to the current level of 11% (in 2017). This suggests that the impact of oligopoly on surplus creation has increased over time.

Figure 3: Total Surplus of US public firms (1997-2017)

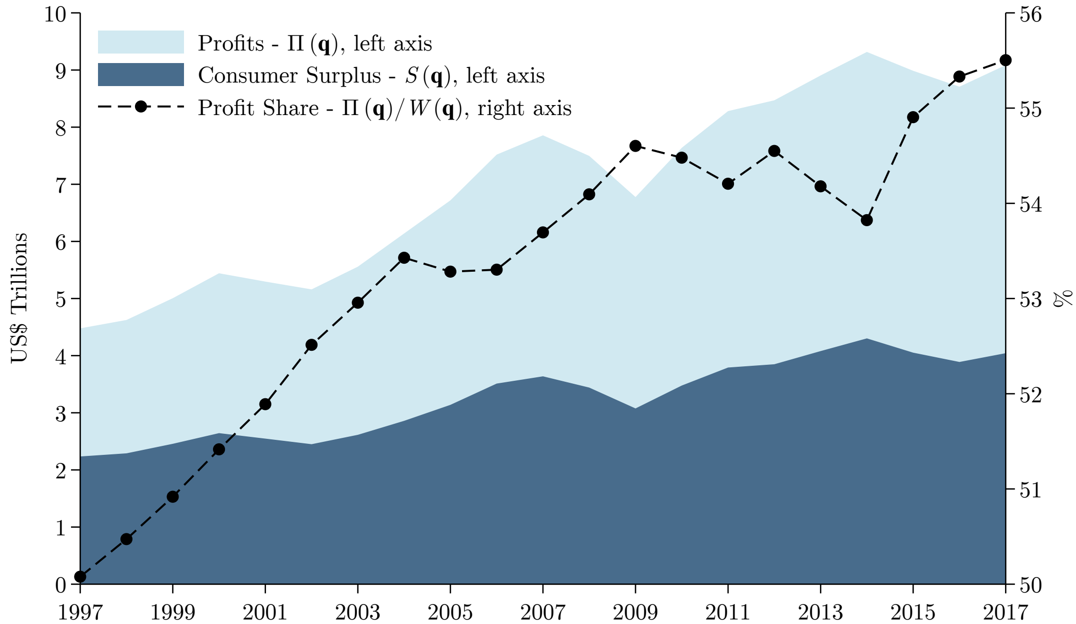


Figure Notes: The figure above plots the evolution, between 1997 and 2017, of aggregate (economic) profits $\Pi(\mathbf{q})$, aggregate consumer surplus $S(\mathbf{q})$ and total surplus $W(\mathbf{q})$, as defined in the model from Section 2. Profits as a percentage of total surplus (Π/W, black dotted line) are shown on the right axis. These statistics are estimated over the universe of the US publicly-listed corporations. These surplus measures are gross of fixed costs. Appendix E replicates this graph using surplus measures that are net of fixed costs.

To investigate the impact of fixed costs on these results, I plot, in the same figure (light line), the percentage difference in total surplus between the Cournot equilibrium and the *Second Best*: it has increased from 6.4 percentage points (in 1997) to 9.3 percentage points (in 2017). In other words, when fixed costs are taken into account, the welfare loss caused by oligopoly starts from a lower level (as is to be expected, mathematically) but increases more sharply over time (+45% over the period)

Overall, my findings are consistent with the interpretation that U.S. public firms have more oligopoly power than they had in 1997, and that this increase in oligopoly power has had a significant impact on both allocative efficiency and consumer welfare.

4.3. The Role of Entry Costs

The increasing concentration of US industries is reflected in the stark decline of the number of public companies: their number has dropped from about 7,500 in 1997 to about 3,500 in 2017 (Kahle and Stulz, 2017). This is largely the result of a collapse in the rate of new Initial Public Offerings (IPOs). This lower rate of entry among public fims, as well as the higher profits and deadweight losses, could be the result of higher entry costs. Obviously, this is hard to know for sure, since entry costs are not directly observable.

Next, I use my model to obtain an estimate of how costs of entry have changed over time. The intuition behind this exercise is to ask how high would the cost of entry have to be, in order to deter a marginal startup from entering the market. Thus far, I have kept the number of granular firms fixed. Now, I consider

Figure 4: Deadweight Loss from Oligopoly (1997-2017)

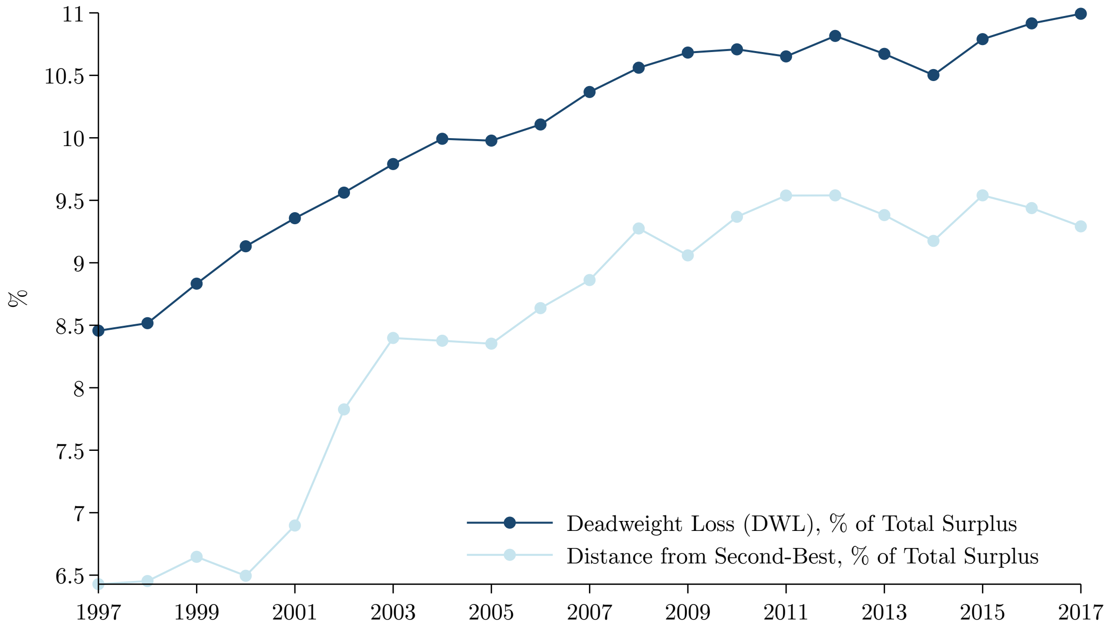


Figure Notes: The following figure plots the estimated deadweight loss (DWL) from oligopoly, between 1997 and 2017. The darker line is the traditionally-defined DWL - the % difference in total surplus between the Cournot equilibrium and the First-Best scenario, while the lighter line is the % difference between the Cournot equilibrium and the Second-Best scenario as defined in Section 2. These surplus measures are gross of fixed costs. Appendix E replicates this graph using surplus measures that are net of fixed costs.

the entry problem of a marginal startup indexed by $i = 0$ that has quality-adjusted productivity $(b_0 - c_0)$ and fixed costs (f_0) that are equal to the median in the population of incumbents.

We can use the model to compute firm 0's economic profits conditional on entry (π_0). Let us then assume that firm 0 faces an entry cost of e_0, and that it enters the market if $\pi_0 - f_0 \ \geq \ e_0$. Given these assumptions, the net profits conditional on entry $(\pi_0 - f_0)$ provide a lower bound for the entry cost e_0.

The challenge in computing π_0 is that, in my model, firms exist in a space of product characteristics. Therefore, to estimate π_0, I also need to measure 0's cosine similarity vis-à-vis every other firm in the dataset $(\mathbf{a}_0'\mathbf{a}_i)$. To measure $\mathbf{a}_0'\mathbf{a}_i$ (for all i), I obtain another set of cosine similarities from Hoberg, Phillips and Prabhala (2014). These cosine similarities are computed, for Compustat firms, against a generic venture capital-backed startup. To construct them, the authors use startup product descriptions from the VenturExpert database. This data is available for 1997-2008.[16]

Figure 5 presents the resulting estimate of the implied entry cost, normalized at the level of 1997. As of 2008, it has increased by approximately 60%. What this exercise tells us is that incentives to enter have increased; hence, from a marginal analysis perspective, the model suggests that constraints or disincentives to enter should also have increased over this period. This occurs at approximately the same time as the number of new IPOs collapses.

[16] I thank the authors for retrieving and sharing this data, which is not part of the public HP repository.

Figure 5: Implied Cost of Entry for a VC-backed Startup

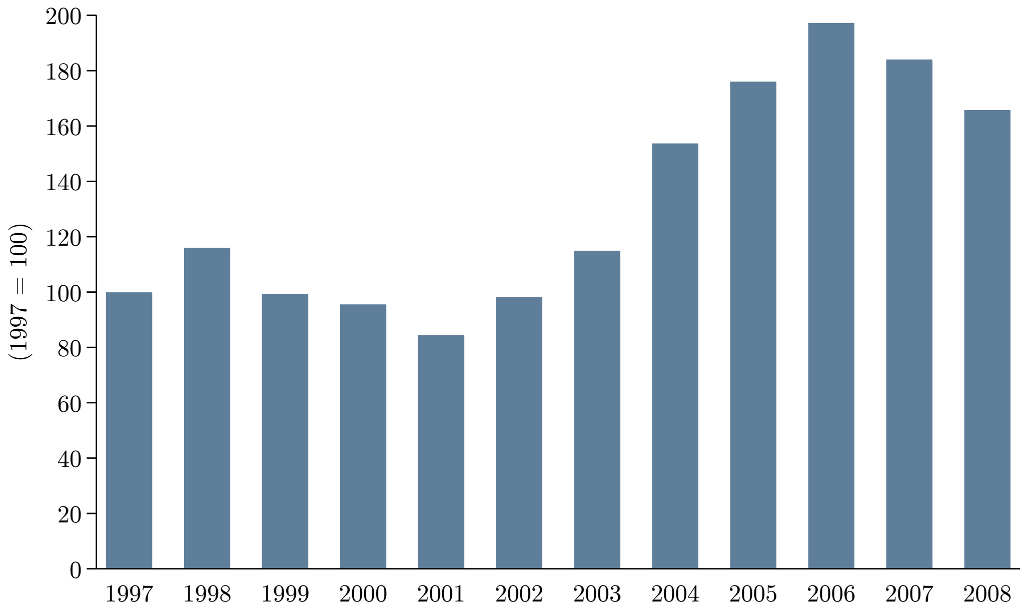


Figure Notes: the figure above plots the implied cost of entry for a VC-backed startup with median quality-adjusted productivity ($b_i - c_i$) and fixed cost (f_i), as implied by the model in Section 2. The similarity scores for the marginal entrant are computed by Hoberg, Phillips and Prabhala (2014), using VenturExpert product descriptions.

This exercise is, obviously, not informative about the nature of these unobserved entry costs. In the next Subsection, I further investigate the collapse of the rate of entry among public firms.

4.4. Startup Takeovers and the Secular Decline of IPOs

One interesting aspect of the decline of IPOs is that it appears to be unrelated to the decline of the startup rate that has been measured in the broader economy (Decker, Haltiwanger, Jarmin and Miranda, 2014). Far from declining, the number of startups that are backed by Venture Capital (VC), which make up the majority of startups that eventually become public companies, has boomed over this period.

Figure 6 plots the number of Venture Capital exits in the United States by year and type, for the period 1985-2017[17]. In the diagram, I separate IPOs from acquisitions. We can see from this figure that, at the beginning of the 1990s, the vast majority of VC exits were IPOs. However, since the mid-90s, there has been a dramatic shift toward acquisitions. One implication of this fact is that the decline of IPOs was not driven by a dearth of startups. Instead, the reason why IPOs have decreased is that they have been largely replaced by acquisitions.

The next question I ask is to what extent can the shift from IPOs to acquisitions account for the rising profit share of surplus, and the increasing deadweight loss?

To investigate this question, I construct a counterfactual in which I add granular firms to the model,

[17]This data is sourced from the National Venture Capital Association (NVCA).

Figure 6: Venture Capital Startup Exits by Type

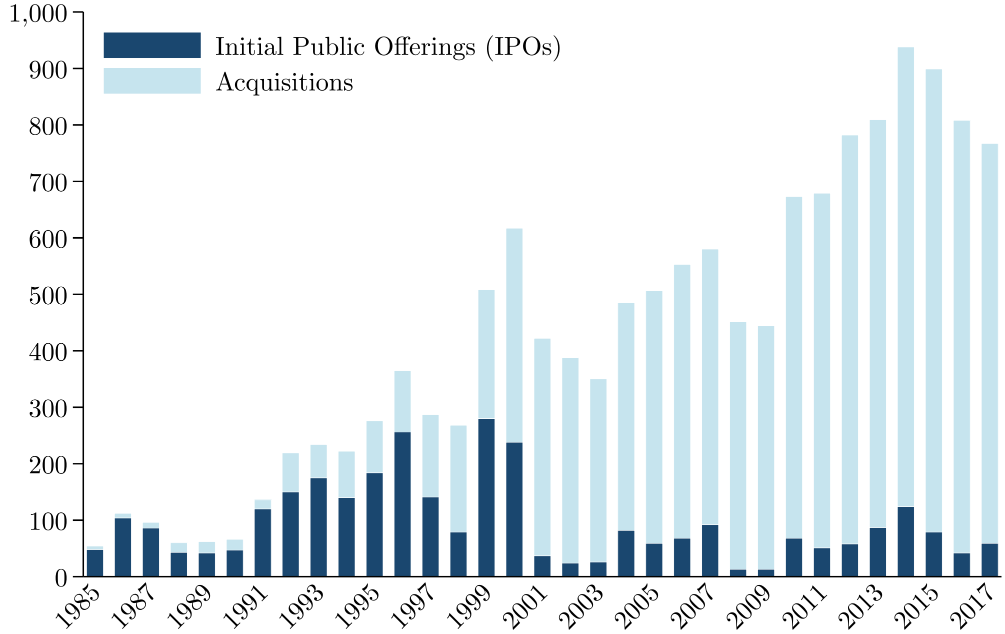


Figure Notes: the figure above plots the number of successful venture capital exits in the United States by year and type (Initial Public Offering v/s Acquisition). The data is sourced from the National Venture Capital Association (NVCA).

with the objective of simulating the ratio of IPOs to acquisitions remaining constant after 1997. For each firm i entering Compustat after 1997, I add $(N_i - 1)$ firms to the model. These firms are "similar" to i in the sense that they share the same value of $(b_i - c_i)$ as well as the same coordinates in the space of common characteristics $(\mathbf{a}_i)$; they also exit the sample in whichever year firm i exits the dataset. Yet, they are not perfect substitutes to i, due to the presence of idiosyncratic characteristics.

N_i is determined so that, in this counterfactual, the ratio of IPOs to acquisitions remains constant after 1997. Specifically, let IR_t be ratio of IPOs to total VC exits at time t. I set N_i as follows:

$$N_i = \begin{cases} \frac{\text{IR}_{1997}}{\text{IR}_t} & \text{if } i \text{ went public at time } t \\ 1 & \text{otherwise} \end{cases} \tag{4.1}$$

We can see the result of this counterfactual analysis in Figure 7: it shows the percent difference in consumer surplus between the Cournot equilibrium and the *Perfect Competition* counterfactual, under two alternative scenarios. The lighter line shows the baseline case: consistent with the findings of Subsection 4.2, the percentage gap in consumer surplus increases from 42% to 51%, reflecting the larger deadweight loss as well as the larger share of total surplus accruing to producers.

The darker line shows the counterfactual where the ratio of IPOs to acquisitions stays constant after 1997. Under this alternative scenario, the increase in the consumer surplus "gap" is significantly less pronounced, leveling at 43.9% in 2017. This reflects a more muted increase in the deadweight loss, as well as a slight decrease in the profit share of surplus.

Figure 7: Consumer Surplus, % difference from Perfect Competition

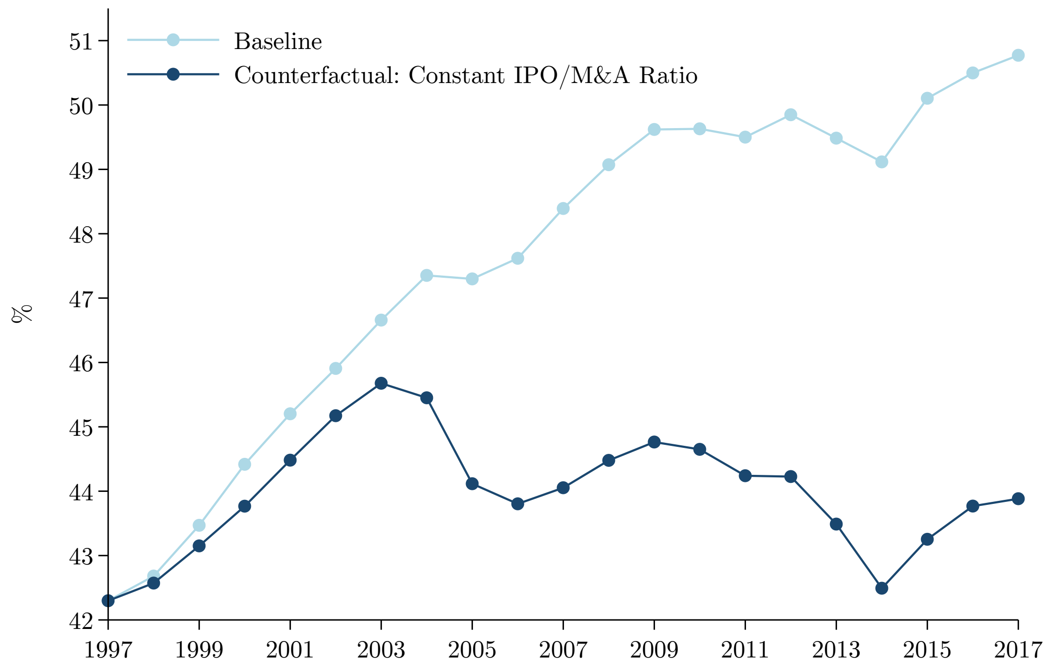


Figure Notes: The following figure plots the percentage difference in consumer surplus between oligopolistic competition and perfect competition, between 1997 and 2017. The observed equilibrium value (light line) is plotted against a counterfactual scenario (darker line) in which the ratio of IPOs to startup acquisitions remains constant after 1997.

This counterfactual analysis suggests that the shift from IPOs to takeovers may have contributed at least to some degree to the measured increase in oligopoly power. There are some caveats, of course. First, this is just one of many factors that may have contributed – and it only offers a *proximate* explanation for the increasing oligopoly power: my model is silent on the reasons why this shift from IPOs to acquisitions has occurred in the first place[18]. It is also important to note that this counterfactual analysis should not be interpreted as causal evidence.

That being said, these results do seem to align with a recent empirical literature that has focused on the anti-competitive effects of startup acquisitions. Wollmann (2019), for example, argues that startup acquisitions may have been used by large corporations to engage in what he calls *stealth consolidation*. The idea is that the majority of startup acquisitions fall under the size threshold for mandatory merger review: as a consequence, startup acquisitions rarely undergo merger review. Large companies may then be able to use startup acquisitions to engage in monopolization with little risk of attracting antitrust scrutiny.

In another study, Cunningham, Ederer and Ma (2018) provide evidence from the pharmaceutical industry of so-called *killer acquisitions*: specifically, they show that a significant share of startup acquisitions by drugmakers results in the arrested development of startups' drugs that might compete with the acquirers' own.

[18]See Bowen, Frésard and Hoberg (2018) and Ewens and Farre-Mensa (2020) for a discussion of the potential causes.

5. Robustness and Extensions

In this section, I investigate the robustness of my empirical analyses to a variety of assumptions; I propose extensions to the model; and I discuss the limitations of this paper and potential follow-up work.

5.1. Private and Foreign Firms, Endogenous Entry

I verify that my baseline results are not sensitive to the inclusion of private and foreign firms. Based on the aggregation result derived in Subsection 2.6, I am able to include non-Compustat companies in the model by adding a representative firm that acts competitively and whose size reflects the endogenous entry and exit of atomistic players. This representative firm is used to capture the economic activity of private and foreign firms.

In order to implement this version of the model, I need to assume a cosine similarity between this representative firm, which I label $n+1$, and every other firm $i = 1, 2, ..., n$. I assume that the cosine similarity between i and $n+1$ is simply equal to the average cosine similarity between i and every other firm – formally:

$$\mathbf{a}_i' \mathbf{a}_{n+1} = \frac{1}{n-1} \sum_{j \neq i} \mathbf{a}_i' \mathbf{a}_j \tag{5.1}$$

My empirical results only change slightly as a consequence of this modification: corporate profits, as a percentage of the surplus produced by public firms, increase from 49.5% in 1997 to 55% in 2017. The deadweight loss increases is virtually unchanged.

To ensure that the approach of modeling non-Compustat firms as atomistic is valid, I investigate whether Compustat becomes a much larger or smaller share of US GDP over time. I compute an estimate of the value added by Compustat companies, and investigate how it changes over time as a percentage of corporate business GDP (obtained from BLS NIPA tables). Reassuringly, I find that this percentage does not trend (either positively or negatively) over the period considered: it is 46.8 in 1997 and 47.4 in 2017, with a standard deviation of 5.9 percentage points over the period.

5.2. Fixed Costs

Thus far, I have defined aggregate profits (Π) and total surplus (W) *gross* of fixed costs (F). Next, I study how Figures 3 and 4 change if we subtract F from Π and from W. In other words, I want to investigate whether the higher economic profits are somehow justified by higher fixed costs.

In Appendix E, I reproduce both these figures, after redefining Π and W to be computed net of fixed costs. By comparing Figures 3 and 4 with Figures 13 and 14, we can see that my core empirical results are largely unaffected by how fixed costs are incorporated in the analysis. The most significant change in the findings is that the profit share of surplus increases much more dramatically when fixed costs are subtracted: from 11% in 1997 to 22% in 2017.

5.3. Intangible Capital

With regard to the estimation of fixed costs, there has been some debate in the literature about how Selling, General & Administrative (SGA) costs should be treated from an accounting standpoint. This item, as

presented in Compustat, includes miscellaneous costs that are not directly linked to production (see Traina, 2018). It also includes R&D expenditures.

While it is generally understood that these are not variable costs, it is also not entirely clear that this cost item is simply overhead. Eisfeldt and Papanikolaou (2013) have argued that SGA partly embeds investments in *intangible capital*, and therefore should not be treated as overhead but capitalized.

Based on this argument, Peters and Taylor (2017) have developed a measure of intangible capital for Compustat: they treat R&D expenditures, plus 30% of the remaining portion of SGA, as investment in intangible capital. They then computed the firm-level intangible capital stock by applying a perpetual inventory model. If we choose to capitalize, rather than expense, these putative investments in intangible capital, we then obtain the following alternative measure of fixed costs:

$$f_i = (\mathrm{SGA}_i - \mathrm{R\&D}_i) \times 0.7 + (\mathrm{PPE}_i + \mathrm{Intangible\,Capital}_i) \times \mathrm{User\,Cost\,of\,Capital} \tag{5.2}$$

Changing the definition of fixed costs (by definition) does not affect my measurements in Figure 3, nor the deadweight loss from Figure 4, because fixed costs do not enter these measures. It does, however, affect the distance from the second-best (which is shown in Figure 4) as well as the analysis from Appendix E, which I have discussed above. In the same Appendix, I show that my results are largely unaffected if intangible investments are capitalized, rather than expensed.

5.4. Labor Supply Elasticity

Next, I investigate how my empirical results change if I assume a completely-inelastic labor supply function.

By definition, profit as a share of total surplus would be unchanged. The deadweight loss would instead become the total surplus difference between the Cournot equilibrium and the *Resource-Efficient* counterfactual, which we previously defined in Subsection 2.5. As can be seen in Table 1, this welfare difference is smaller than the deadweight loss. Intuitively, this is because the labor supply (by definition) cannot respond to the removal of the oligopolistic distortions.

I compute this alternative measure of the deadweight loss (the percentage difference in total surplus between Cournot and *Resource Efficient*) over the period 1997-2017. I find that my core empirical results carry through: the level of this "alternative" deadweight loss is 5.2% in 1997, and it increases to 7.9% by 2017. In other words, the level of the deadweight loss is lower if we assume a fixed labor supply (as should be expected), but it increases more sharply (by half) over the 20-year period.

5.5. Multi-product Firms (Diversification vs. Differentiation)

Like most of the macroeconomics literature, this paper does not use product-level data on characteristics. Hence, while the model does reasonably well in capturing measured variation in markups across firms (see Appendix D), it cannot speak to how the market power of a firm varies across individual product markets (say, how Apple's market power in the smartphone industry may be higher or lower than in the personal computer market). Nevertheless, given the extensive presence in Compustat of multi-product firms, it is worth investigating what additional assumptions are required for my model to apply to multi-product firms.

Suppose that there are still n firms and k characteristics, but now the n firms produce a total of $m \geq n$ products. The same product might be produced by multiple firms and the same firm may produce more

than one product. The vector of units produced for each good is now the m-dimensional vector $\mathbf{y}$. Similarly to matrix $\mathbf{A}$ in Section 2, matrix $\mathbf{A}_1$ transforms units of products into units of characteristics:

$$\mathbf{x} = \mathbf{A}_1\mathbf{y} \tag{5.3}$$

Because firms are diversified, each firm now produces a basket of goods: instead of representing the number of units produced of each product, the vector $\mathbf{q}$ now represents the number of baskets produced by each firm. The matrix $\mathbf{A}_2$ transforms the vector of quantity indices into units of products supplied:

$$\mathbf{y} = \mathbf{A}_2\mathbf{q} \tag{5.4}$$

Now I put together the previous two equations. Letting $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2$, I have

$$\mathbf{x} = \mathbf{A}_1\mathbf{y} = \mathbf{A}_1\mathbf{A}_2\mathbf{q} = \mathbf{A}\mathbf{q} \tag{5.5}$$

The identity above demonstrates how the linear-hedonic structure makes the model easily generalizable to multi-product firms. The intuition is that, if the output of a firm i is not a single product, but rather a basket of products, one can project the quantity index q_i on the characteristics space either in two steps (by projecting it first on goods and then on characteristics), or in one single step (using the composite projection matrix $\mathbf{A}$).

The limitation of this multi-product interpretation of the model is that, while firms can change their supply q_i, the vector $\mathbf{a}_i$ must stay fixed. What this means is that while firms may produce more than one product and can scale up or down the quantity of baskets produced, they must keep producing the products in constant quantity ratios.

A different way of saying this is that the limitation of firm-level data is that it does not allow to study the reallocation of resources *within* firms, but only *across* firms. This implies that my estimates of the deadweight loss may be conservative: if firms have different degree of market power in different markets, this will generate *within-firm* variation in markups. My model does not capture the additional welfare gains that would be realized if the social planner could remove *within-firm* dispersion in markups.

It is important to emphasize that this limitation is not specific to this paper, but it is endemic in the literature. While my linear demand specification does not fully address it, I claim that the GHL demand system handles multi-product firms better than CES preferences.

Additionally, suppose that we had product or plant-level data (including similarity scores for individual products), so that we may relax the assumption that $\mathbf{A}_2$ is fixed (firms can change output ratios among products). The model would then need to account for the fact that different business units within the same firm coordinate their supply decisions, similar to what merging/colluding firms do in the counterfactual from equation (2.38). Hence, one argument for using firm-level data is that it is a tractable way of modeling coordination of plants within the same firm.

Finally, it useful to ask how does diversification affect the measured intensity of competition across firms. If, for example, firms became more diversified, would the model erroneously interpret the resulting change in cosine similarity as an increase in market power? The answer to this question is *it depends* – i.e. the effect of diversification of measured competition is ambiguous. It is possible to construct examples where diversification leads to higher, lower, or unchanged cosine similarity. In Appendix I, I construct three such

examples to illustrate this argument.

5.6. Complements

Because the matrix $\mathbf{\Sigma}$ is non-negative by construction, the marginal utility from one unit of product j is always non-increasing in q_i - formally:

$$\frac{\partial^2 S}{\partial q_i \partial q_j} \;=\; -\sigma_{ij} \;\leq\; 0 \quad \forall\, i \neq j \tag{5.6}$$

In light of equation (5.6), it is tempting to jump to the conclusion that all products are by construction substitutes and that no pair of products are complements. That conclusion is, however, incorrect.

To understand why, we need to recall the textbook definition of substitution and complementarity. Two goods (i, j) are[19]:

$$\text{Complements} \quad \text{if} \quad \frac{\partial q_i}{\partial p_j} < 0 \qquad\qquad \text{Substitutes} \quad \text{if} \quad \frac{\partial q_i}{\partial p_j} > 0 \tag{5.7}$$

We intuitively expect this derivative to have the opposite sign of that in equation (5.6). In the case of CES, this intuition is correct. In the case of my model, however, this intuition fails. This is a consequence of the fact that the cross-price demand elasticity depends on the inverted matrix $(\mathbf{I}+\mathbf{\Sigma})^{-1}$, not on $\mathbf{\Sigma}$ itself. If $\mathbf{\Sigma}$ is not symmetric (here it is not) the off-diagonal elements of $-(\mathbf{I}+\mathbf{\Sigma})^{-1}$ will generally include positive as well as negative elements.[20] This implies that, in the empirical implementation of the model, many producer pairs are strategic complements.

For example, if we compute the vector of cross-price derivatives for car manufacturer General Motors in 2017, we will find that it includes several negative elements (i.e. complements), mostly corresponding to energy and consumer finance companies. This makes sense: intuitively, we expect higher oil prices, loan rates or insurance premia to adversely affect the residual demand for cars.

Hence, despite the property of the model described by equation (5.6), my model does produce strategic complementarity. Indeed, I argue that one of the strengths of the network Cournot model is its ability to produce a rich competitive environment that includes complement goods.

5.7. Limitations and Future Work

This model (like every other model) has limitations and leaves out certain aspects of market power that might be relevant to the current debate on antitrust policy.

For example, one important assumption that I make in my model (in order to preserve tractability) is that all firms are final goods firms. In other words, input-output linkages between individual firms are not part of the model. This might lead to underestimating the welfare costs of oligopoly, if input-output linkages result in double marginalization. One interesting extension of the model would be to modify the firm's production

---

[19] Usually, these derivatives refer to the *Hicksian* demand, to ensure that $\operatorname{sign}(\partial q_i/\partial p_j) = \operatorname{sign}(\partial q_j/\partial p_i)$ (see Kreps, 2012, section 11.6). This distinction can be ignored for my model, since the sign of this derivative is the same for the Marshallian and the Hicksian demand.

[20] It is fairly easy to come up with examples. Consider for instance three goods (1,2,3) and three common characteristics (A,B,C). If goods 1 and 3 load entirely on characteristic A and C respectively, and good 2 loads equally on all three characteristics, then it can be verified that goods 1 and 3 are strategic complements.

function to allow firms to use other firm's output as inputs. In order to take this extended model to the data, granular input-output data on firm-to-firm relationships would be required.

Another important restriction of my model is that it treats the firms' position in the product characteristics space as fixed. The assumption that product characteristics are exogenous is standard in the demand estimation literature (Berry et al., 1995; Nevo, 2001); however, it makes policy simulations less robust in the long-run. This is because, given enough time, firms may be able to endogenously change their product portfolios. Some recent work in the IO literature (see Fan, 2013; Wollmann, 2018) has considered endogenous product characteristics. If product characteristics data can be obtained (HP only provide cosine similarities), another interesting direction for future research is to generalize my model by endogenizing the firms' position in the characteristics space.

Finally, another force that is left out of this paper is labor market power. Concentration may lead not only to oligopoly power in output markets: it may also lead to oligopsony in input markets. Because the model that I presented does not speak to this channel, my estimates of income distribution and aggregate efficiency do not include the effects of the labor market power, which might also have increased over this period.

With additional labor market data, it is definitely possible to apply the methods developed in this paper to a labor market setting. Cosine similarities can easily be constructed for textual descriptions of job vacancies: if we see labor input as a differentiated good, we can then develop and estimate a labor market version of the model presented in Section 2. In Appendix J, I propose two ways of reframing my model to study labor market power: the first considers *workers'* monopoly power (workers with unique characteristics command higher wages); the latter looks at oligopsony (firms that utilize unique inputs are able to charge higher markups).

6. Conclusions

In this study, I have presented a new general equilibrium model of oligopolistic competition with hedonic demand and differentiated products, with the objective of measuring the welfare consequences of rising oligopoly in the United States from 1997 to 2017. To estimate my model, I used a data set (recently developed by Hoberg and Phillips, 2016) of product cosine similarities that covers all public firms in the United States on a yearly basis. Through the lens of my model, these similarity scores are used to retrieve the cross-price elasticity of demand for every pair of publicly traded firms.

My measurements suggest that oligopoly has a considerable and growing effect on aggregate welfare. In particular, I estimate that, if all publicly traded firms were to behave as atomistic competitors, the total surplus produced by this set of companies would increase by 11 percentage points. Consumer welfare would increase even more dramatically—it would more than double—as surplus would be largely reallocated from producers to consumers. I find that a large share of the deadweight loss caused by oligopoly (7.7 percentage points) can be attributed to resource misallocation—that is, a significant share of the deadweight losses could theoretically be recovered by a benevolent social planner, even if we assumed labor to be inelastically supplied.

I also find evidence that the deadweight losses increases significantly in the presence of collusion: consolidating firm ownership in the hands of one owner, who induces firms to coordinate, would depress aggregate surplus by an additional 10 percentage points. Consumer surplus would drop even more, with a projected

decrease of about 38 percentage points. Overall, my analysis of firm-level data suggests that there is evidence of sizable welfare distortions due to oligopoly power.

By mapping my model to firm-level data for every year between 1997 and 2017, I find that, while both the profits earned by U.S. public corporations and the corresponding consumer surplus have increased over this period, profits have increased at a significantly faster pace: consequently, the share of surplus appropriated by firms in the form of profits has increased substantially (from 50% to 55.5%). Consistent with this finding, I estimate that the welfare costs of oligopoly, computed as the percentage increase in surplus that is obtained by moving to the competitive outcome, have increased (from 8.5% to 11%). Overall, my estimates are consistent with the hypothesis that the observed secular trends in markups and concentration have resulted in an increased welfare loss, particularly at the expense of the consumer.

The model allows me to compute a number of novel counterfactuals that are relevant to competition policy, and to shed additional light on these ongoing trends. I have shown that a potential contributor to the increased oligopoly power might lie in the secular decline in IPOs and the surge in takeovers of VC-backed startups. Through the lens of my model, this shift can account for a large share of the increase in the deadweight loss and in the producer share of surplus.

This paper contributes—both methodologically and empirically—to a growing literature in macroeconomics and finance that is devoted to incorporating heterogeneity, imperfect competition and Industrial Organization methods in general equilibrium models. In particular, it shows that combining firm financials with measures of similarity based on natural-language processing of regulatory filings offers a promising avenue to model product differentiation and imperfect substitutability at the macroeconomic level: it affords the opportunity to impose a less arbitrary structure on the degree of substitution across sectors and firms.

In addition to the theory contribution, this paper provides measurements that align with a growing body of empirical work on rising market power (De Loecker, Eeckhout and Unger, 2020) and the anti-competitive effects of startup acquisitions (Cunningham et al., 2018; Wollmann, 2019). One potential policy implication of my findings is that startup takeovers may have significant anti-competitive effects at the aggregate level – even if one takes the view that individual acquisitions have negligible impact. Because competition authorities are much less likely to challenge these acquisitions than they are to oppose a merger between large incumbents, these findings strengthen the case for increasing the antitrust oversight of these transactions.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The network origins of aggregate fluctuations," *Econometrica*, 80, 1977–2016.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2017): "Microeconomic origins of macroeconomic tail risks," *American Economic Review*, 107, 54–108.

Asplund, M., and V. Nocke (2006): "Firm turnover in imperfectly competitive markets," *The Review of Economic Studies*, 73, 295–327.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020): "The fall of the labor share and the rise of superstar firms," *The Quarterly Journal of Economics*, 135, 645–709.

Axtell, R. L. (2001): "Zipf Distribution of US Firm Sizes," *Science*, 1818–1820.

Baker, J. B., and T. F. Bresnahan (1985): "The gains from merger or collusion in product-differentiated industries," *The Journal of Industrial Economics*, 427–444.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006): "Who's who in networks. Wanted: The key player," *Econometrica*, 74, 1403–1417.

Baqaee, D. R., and E. Farhi (2020): "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 841–890.

Bloom, N., M. Schankerman, and J. Van Reenen (2013): "Identifying technology spillovers and product market rivalry," *Econometrica*, 81, 1347–1393.

Bonacich, P. (1987): "Power and centrality: A family of measures," *American journal of sociology*, 92, 1170–1182.

Bowen, D. E. I., L. Frésard, and G. Hoberg (2018): "Technological disruptiveness and the evolution of IPOs and sell-outs."

Carvalho, V. M. (2014): "From micro to macro via production networks," *Journal of Economic Perspectives*, 28, 23–48.

Carvalho, V. M., M. Nirei, Y. Saito, and A. Tahbaz-Salehi (2020): "Supply chain disruptions: Evidence from the great east japan earthquake," *The Quarterly Journal of Economics, forthcoming.*

Carvalho, V. M., and A. Tahbaz-Salehi (2019): "Production networks: A primer," *Annual Review of Economics*, 11, 635–663.

Chen, H., L. Cohen, and D. Lou (2016): "Industry window dressing," *The Review of Financial Studies*, 29, 3354–3393.

Chintagunta, P. K. (2002): "Investigating category pricing behavior at a retail chain," *Journal of Marketing Research*, 39, 141–154.

Covarrubias, M., G. Gutiérrez, and T. Philippon (2020): "From Good to Bad Concentration? US Industries over the past 30 years," *NBER Macroeconomics Annual*, 34, 1–46.

Cunningham, C., F. Ederer, and S. Ma (2018): "Killer acquisitions," *Available at SSRN 3241707*.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*.

De Loecker, J., and P. T. Scott (2016): "Estimating market power: Evidence from the US Brewing Industry," working paper, National Bureau of Economic Research.

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014): "The role of entrepreneurship in US job creation and economic dynamism," *Journal of Economic Perspectives*, 28, 3–24.

Dhingra, S., and J. Morrow (2019): "Monopolistic competition and optimum product diversity under firm heterogeneity," *Journal of Political Economy*, 127, 196–232.

Edmond, C., V. Midrigan, and D. Y. Xu (2018): "How costly are markups?" working paper, National Bureau of Economic Research.

Einav, L., and J. Levin (2010): "Empirical industrial organization: A progress report," *Journal of Economic Perspectives*, 24, 145–62.

Eisfeldt, A. L., and D. Papanikolaou (2013): "Organization capital and the cross-section of expected returns," *The Journal of Finance*, 68, 1365–1406.

Epple, D. (1987): "Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products," *Journal of political economy*, 95, 59–80.

Ewens, M., and J. Farre-Mensa (2020): "The Deregulation of the Private Equity Markets and the Decline in IPOs," *The Review of Financial Studies*, 10.1093/rfs/hhaa053, hhaa053.

Fan, Y. (2013): "Ownership consolidation and product characteristics: A study of the US daily newspaper market," *American Economic Review*, 103, 1598–1628.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?" *American Economic Review*, 98, 394–425.

Fruchterman, T. M., and E. M. Reingold (1991): "Graph drawing by force-directed placement," *Software: Practice and experience*, 21, 1129–1164.

Gao, X., J. R. Ritter, and Z. Zhu (2013): "Where have all the IPOs gone?" *Journal of Financial and Quantitative Analysis*, 48, 1663–1692.

Goeree, M. S. (2008): "Limited information and advertising in the US personal computer industry," *Econometrica*, 76, 1017–1074.

Grullon, G., Y. Larkin, and R. Michaely (2019): "Are US Industries Becoming More Concentrated?" *Review of Finance*, 23, 697–743.

Hoberg, G., and G. Phillips (2016): "Text-based network industries and endogenous product differentiation," *Journal of Political Economy*, 124, 1423–1465.

Hoberg, G., G. Phillips, and N. Prabhala (2014): "Product market threats, payouts, and financial flexibility," *The Journal of Finance*, 69, 293–324.

Hopenhayn, H. A. (1992): "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1127–1150.

Kahle, K. M., and R. M. Stulz (2017): "Is the US public corporation in trouble?" *Journal of Economic Perspectives*, 31, 67–88.

Katz, L. (1953): "A new status index derived from sociometric analysis," *Psychometrika*, 18, 39–43.

Khan, L. (2018): "The new Brandeis movement: America's antimonopoly debate."

Kreps, D. M. (2012): *Microeconomic Foundations I: Choice and Competitive Markets*: Princeton University Press.

Lancaster, K. J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74, 132–157.

Melitz, M. J., and G. I. Ottaviano (2008): "Market size, trade, and productivity," *The review of economic studies*, 75, 295–316.

Monderer, D., and L. S. Shapley (1996): "Potential games," *Games and economic behavior*, 14, 124–143.

Nevo, A. (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69, 307–342.

Peters, R. H., and L. A. Taylor (2017): "Intangible capital and the investment-q relation," *Journal of Financial Economics*, 123, 251–272.

Rosen, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy*, 82, 34–55.

Syverson, C. (2019): "Macroeconomics and Market Power: Context, Implications, and Open Questions," *Journal of Economic Perspectives*, 33, 23–43.

Traina, J. (2018): "Is aggregate market power increasing? production trends using financial statements," working paper, Chicago Booth Stigler Center.

Ushchev, P., and Y. Zenou (2018): "Price competition in product variety networks," *Games and Economic Behavior*, 110, 226–247.

Wollmann, T. G. (2018): "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles," *American Economic Review*, 108, 1364–1406.

——— (2019): "Stealth consolidation: Evidence from an amendment to the Hart-Scott-Rodino Act," *American Economic Review: Insights*, 1, 77–94.

Online Appendices

Product Differentiation and Oligopoly:
a Network Approach

Bruno Pellegrino, UMD

A. Derivation of the Cournot Potential

In the Network Cournot model, each firm i chooses its own output level to maximize its own profit by taking as given the output of every other firm:

$$q_i^* = \arg\max_{q_i} \pi\left(q_i; \bar{\mathbf{q}}_{-i}\right) \tag{A.1}$$

where $\overline{\mathbf{q}}_{-i}$ is the vector of output for every firm except i. The upper bar sign $\bar{\cdot}$ indicates that i treats the quantity supplied by other firms as fixed. The system of first order conditions for this problem is

$$0 = \; b_i - c_i - (2 + \delta_i)\, q_i - \alpha \sum_{j \neq i} \left(\mathbf{a}_i' \mathbf{a}_j\right) \bar{q}_j \tag{A.2}$$

which can be expressed, in vector form, as:

$$0 = \mathbf{b} - \mathbf{c} - \left(2\mathbf{I} + \boldsymbol{\Delta}\right) \mathbf{q} - \boldsymbol{\Sigma} \bar{\mathbf{q}} \tag{A.3}$$

This system of reaction functions defines a vector field $\mathbf{q}\left(\bar{\mathbf{q}}\right)$ which represents the firms' best response as a function of every other firms' strategy. To find the Cournot-Nash Equilibrium, we look for the fixed point $\mathbf{q}^*$ such that $\mathbf{q} = \bar{\mathbf{q}} = \mathbf{q}^*$. Plugging this inside the equation above yields the first order condition that is needed to maximize the potential function $\Phi\left(\mathbf{q}\right)$:

$$0 = \left(\mathbf{b} - \mathbf{c}\right) - \left(2\mathbf{I} + \boldsymbol{\Delta} + \boldsymbol{\Sigma}\right) \mathbf{q}^* \tag{A.4}$$

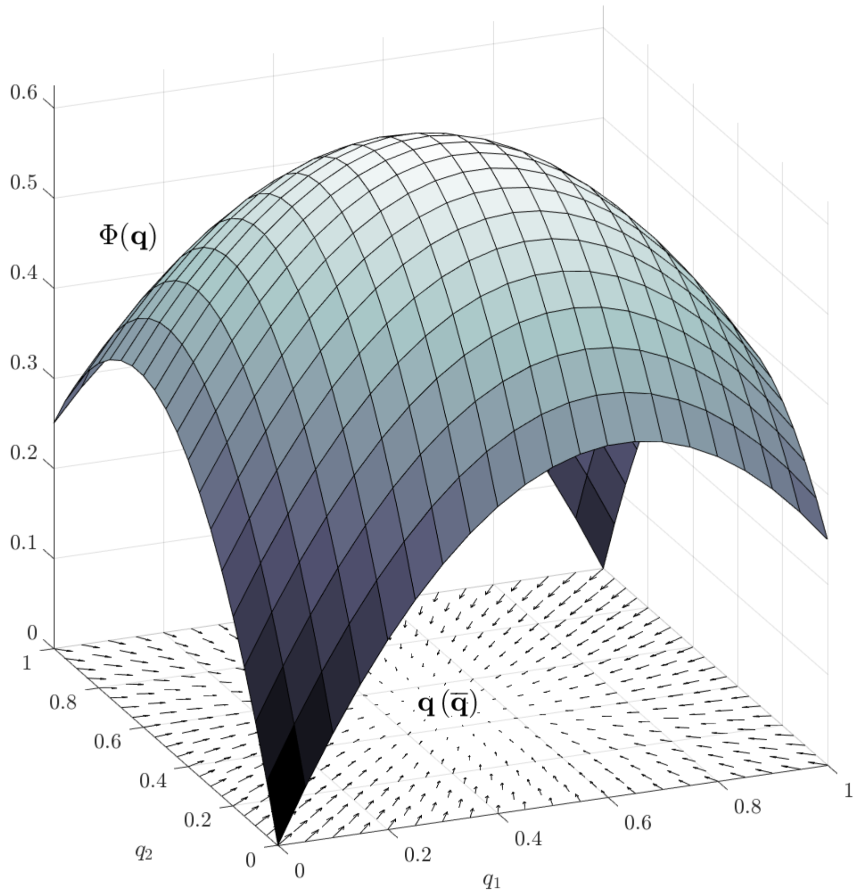
which clarifies why the maximizer of the potential function solves the Network Cournot game. The potential $\Phi\left(\mathbf{q}\right)$ is then obtained as the solution to the following system of partial differential equations

$$\nabla\, \Phi\left(\mathbf{q}\right) \;=\; \left(\mathbf{b} - \mathbf{c}\right) - \left(2\mathbf{I} + \boldsymbol{\Delta} + \boldsymbol{\Sigma}\right) \mathbf{q} \tag{A.5}$$

which equates the gradient of the potential function to the linear system of Cournot reaction functions.

The relationship between the potential and the Cournot-Nash equilibrium is represented graphically, for the two-firm case, in Figure 8. The arrows represent the vector field defined by the firms' reaction functions. The potential function is defined to be the scalar-valued function whose gradient coincides with this vector field. A game is a potential game if the vector field defined by the players' reaction functions is a conservative field - that is, if it is the gradient of some scalar function. We call that function the game's potential.

Figure 8: Graphing the Cournot Potential for the Two-firm Case



B. Nash-Cournot Equilibrium and Network Centrality

In this appendix, I provide additional details on the relationship between equation (2.27), which describes the equilibrium size of firms in my model, and the measures of network centrality developed by Katz (1953) and Bonacich (1987), which are widely used in the social networks literature.

The game played by the firms from Section 2 is a linear quadratic game played over a weighted network. Ballester, Calvó-Armengol and Zenou (2006, henceforth BCZ) show that players' equilibrium actions and payoffs in this class of games depends on their centrality in the network.

In the game played the firms that populate by model, the adjacency matrix of the network over which the game is played, is given by the matrix $(-\mathbf{\Sigma})$. This matrix appears in the quadratic term of all the welfare functions (profits, total surplus and the Cournot potential).

Before discussing how the linkage extends to my model, I am going to formally define the metric of centrality.

Definition 9 (Katz-Bonacich Centrality)**.** For a weighted network with adjacency matrix $\mathbf{G}$, we define the vector of centralities $\mathbf{f}$, with parameters $(\lambda, \mathbf{z})$:

$$\begin{aligned} \mathbf{f}(\mathbf{G}; \lambda, \mathbf{z}) : \qquad \mathbf{f} &= \lambda\, \mathbf{G}\, \mathbf{f} + \mathbf{z} \\ &= (\mathbf{I} - \lambda\, \mathbf{G})^{-1} \mathbf{z} \end{aligned} \tag{B.1}$$

The Katz-Bonacich centrality is defined recursively: a node receives a higher centrality score the higher is the centrality of the nodes it is connected to.

The Nash-Bonacich linkage extends to my model. Suppose there are constant returns to scale ($\mathbf{\Delta} = \mathbf{0}$): then, the Cournot-Nash equilibrium allocation of the model presented in Section 2 (equation 2.27) can be easily verified to coincide with the vector of Katz-Bonacich centralities, with parametrization $\left(\frac{1}{2},\ \frac{1}{2}(\mathbf{b} - \mathbf{c})\right)$:

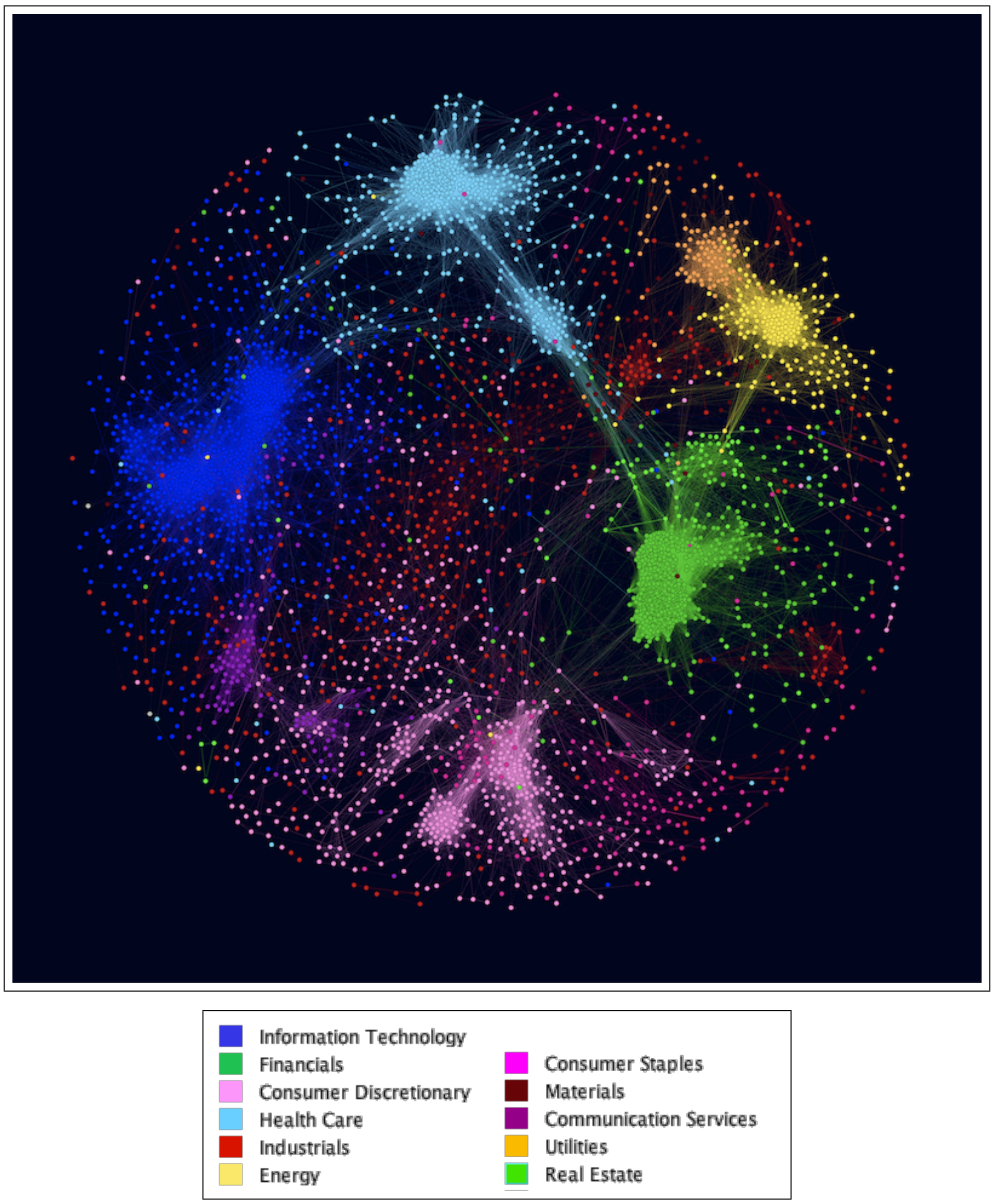
$$\mathbf{q}^{\Phi} \ \equiv\ \mathbf{f}\left(-\mathbf{\Sigma};\ \frac{1}{2},\ \frac{1}{2}(\mathbf{b} - \mathbf{c})\right) \tag{B.2}$$

The peculiarity of the Cournot game played by the firms in my model is that it is played over a *negatively-weighted* network (the adjacency matrix is $-\mathbf{\Sigma}$). The consequence is that the interpretation of $\mathbf{q}^{\Phi}$ as a measure of centrality is reversed: a higher centrality actually reflects a more peripheral position in the network with positive weights $\mathbf{\Sigma}$.

C. Independent Validation of the Hoberg & Phillips Dataset

In this appendix, I validate independently the text-based product similarity measures of Hoberg and Phillips (2016). In the figure below I produce a graph similar to that of Figure 2, while coloring different nodes according to the respective firm's GIC economic sector. The figure shows that there is significant overlap between the macro clusters of the network of product similarity and the broad GIC sectors. To produce this visualization, the dimensionality of data has been reduced from 61,000 to 2; yet, the overlap is nonetheless very clearly visible. The GIC sectors were *not* targeted in producing this graph.

Figure 9: Visualization of the product space (alternate coloring)



D. Data and Calibration

D.1. Mapping to Data and Identification

I present in Table 2 the correspondence between model variables and actual data, with sources. In constructing the dataset, I follow De Loecker, Eeckhout and Unger (2020, henceforth DEU) in excluding firms with negative revenues or costs of goods sold, or negative gross margin (revenues less COGS).

I also follow DEU in the computation of the user cost of capital, which is equal to the federal funds rate (FEDFUNDS from FRED), minus capital goods inflation (PIRIC from FRED), plus a combined depreciation rate and risk premium set at 12%.

D.2. Calibrating the Cost Function

The firms' variable cost function (in linear algebra notation) is:

$$h_i = c_i q_i + \frac{\delta_i}{2} q_i^2 \tag{D.1}$$

We have seen in Section (3) how to recover the intercepts c_i given the slopes δ_i. Calibrating the cost function means calibrating the slopes δ_i. Recall that firms produce using the following quasi-Cobb-Douglas production function:

$$q_i = k_i^{\theta} \cdot \ell(h_i) \tag{D.2}$$

where k_i is the capital input. Following De Loecker, Eeckhout and Unger (2020, henceforth DEU), who also use Compustat data, I measure capital in Compustat as Property, Plant and Equipment). DEU estimate the production function econometrically. I use DEU's estimates to calibrate $\theta = 0.15$.

Then it must be the case that $\ell(\cdot)$ respects:

$$h_i = \ell^{-1}\left(\frac{q_i}{k_i^{\theta}}\right) \stackrel{\text{def}}{=} \tilde{c}_i \left(\frac{q_i}{k_i^{\theta}}\right) + \frac{\tilde{\delta}}{2}\left(\frac{q_i}{k_i^{\theta}}\right)^2 \tag{D.3}$$

where, by definition:

$$c_i = \frac{\tilde{c}_i}{k_i^{\theta}} \qquad \text{and} \qquad \delta_i = \frac{\tilde{\delta}}{k_i^{2\theta}} \tag{D.4}$$

I make the assumption that $\tilde{\delta}$ is constant across firms and over time. Through this assumption, I am reducing the dimensionality of the cost parameter vector to one $\left(\tilde{\delta}\right)$. We can then rewrite equation (3.10) – the markup of firm i – as a function of $\tilde{\delta}$ and observables:

$$\mu_i = \frac{\left(2k_i^{2\theta} + \tilde{\delta}\right) \cdot p_i q_i}{2k_i^{2\theta} \cdot h_i + \tilde{\delta} \cdot p_i q_i} \tag{D.5}$$

where $(p_i q_i, h_i, k_i)$ are all obtained from Compustat.

Table 2: Variable Definitions and Mapping to Compustat

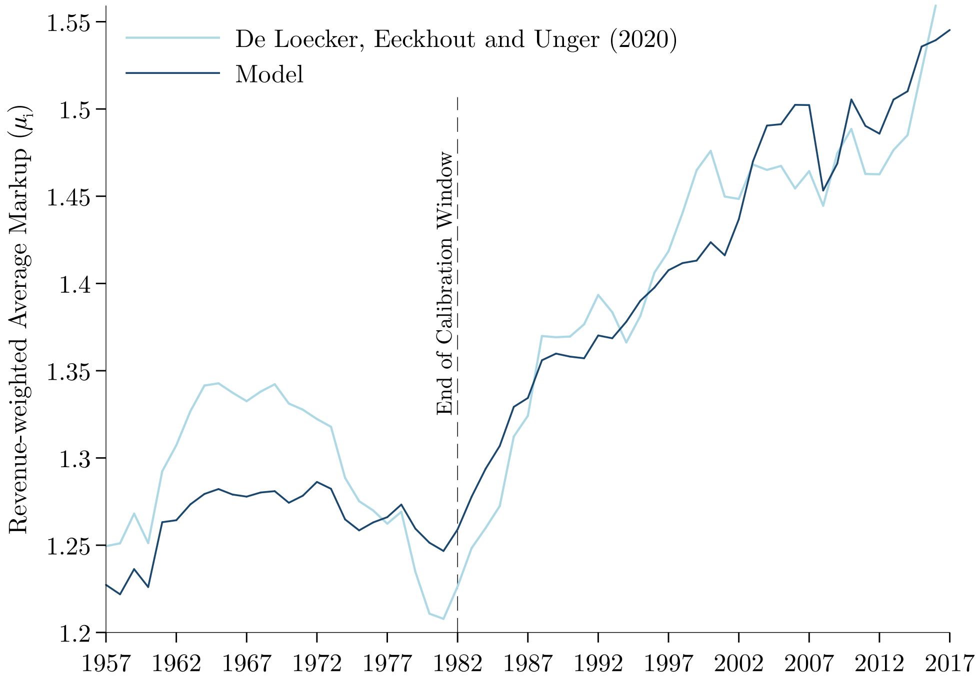
Panel A: Observed Variables

Notation	Concept	Measurement
$p_i q_i$	Revenues	Revenues (*source*: *Compustat*)
h_i	Total Variable Costs	Costs of Goods Sold (*source*: *Compustat*)
f_i	Fixed Costs *alternative measure*:	$\text{SGA}_i + \text{Property Plant \& Equipment}_i \times \text{User Cost of Capital}$ $(\text{SGA}_i - \text{R\&D}_i) \times 0.7 + (\text{Property Plant \& Equipment}_i + \text{Intangible Capital}_i) \times \text{UCC}$ (*source*: *Compustat*, *Federal Reserve Economic Data*)
$\mathbf{a}_i'\mathbf{a}_j$	Product Cosine Similarity	Word cosine similarity in 10-K product description (*source*:*Hoberg and Phillips, 2016*)

Panel B: Identified Variables

Notation	Derived Variable	Computation/Identification
q_i	Output	$= \sqrt{\frac{\pi_i}{1+\delta_i/2}}$
c_i	Marginal Cost Intercept	$= h_i/q_i - \frac{1}{2}\delta_i q_i$
$\mathbf{b}$	Demand Intercept	$= (2\mathbf{I} + \mathbf{\Delta} + \mathbf{\Sigma})\,\mathbf{q} + \mathbf{c}$

Figure 10: Revenue-Weighted Average Markup



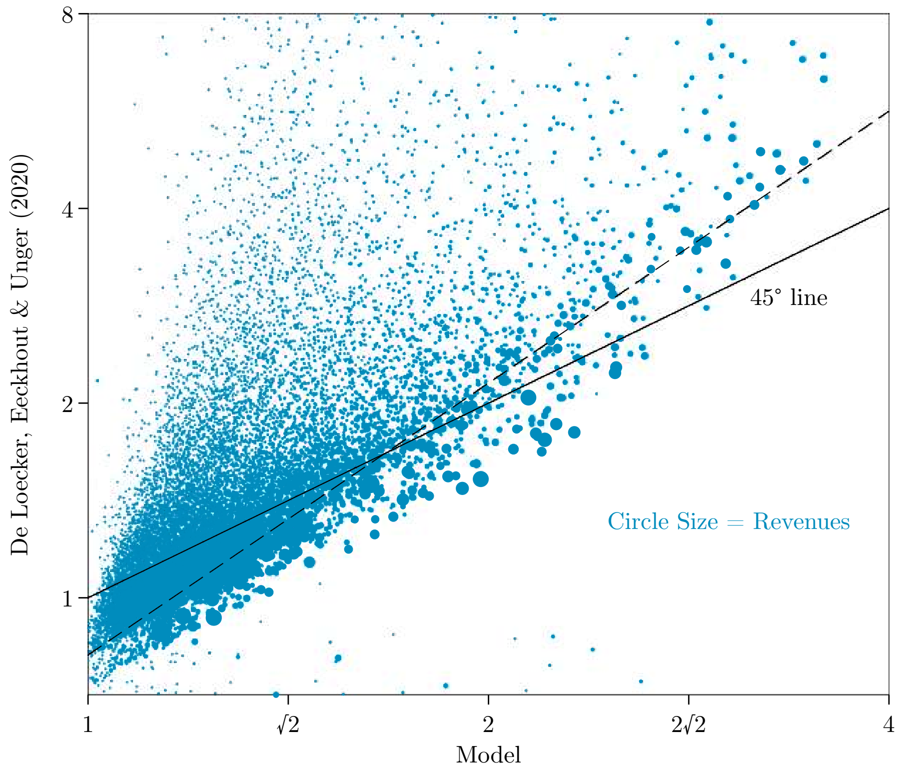
To calibrate $\tilde{\delta}$, I target the markup estimates of De Loecker, Eeckhout and Unger (2020, DEU). Specifically, I calibrate $\tilde{\delta}$ so that the model-inferred average markup matches exactly that estimated by DEU:

$$\tilde{\delta} \; : \quad \mathbb{E}\left[\frac{\left(2k_i^{2\theta}+\tilde{\delta}\right)\cdot p_i q_i}{2k_i^{2\theta}\cdot h_i+\tilde{\delta}\cdot p_i q_i}\right] \; = \; \mathbb{E}\left(\mu_i^{\text{DEU}}\right) \tag{D.6}$$

Following DEU, I weight markups by revenues in order to compute this average. While I can only estimate my full model for the period 1997-2017, I can actually compute markups as far back as the 1950s, since the identification of markups does not require HP's product similarity data.

I use GMM to solve the moment equation (D.6). I pool data from the years preceding the rise in markups that was documented by DEU (1957-1982). This implies that only the long-term average markup before 1982 is targeted (not the time trend). The rationale for this approach is that I can then test the assumption that $\tilde{\delta}$ is stable over time by evaluating whether the $\tilde{\delta}$ calibrated before 1982 allows to capture the rise of markups that is observed after 1982. I obtain $\tilde{\delta} = 6.3$.

Figure 11: Markups: Model vs. DEU Estimates



D.3. Non-Targeted Moments (Model Fit)

For the model to realistically capture the deadweight losses from oligopoly it should generate meaningful time-series and cross-sectional variation in markups. Moreover, we would expect the model-implied markups to correlate with existing econometric estimates (for the Compustat universe, the benchmark would be DEU).

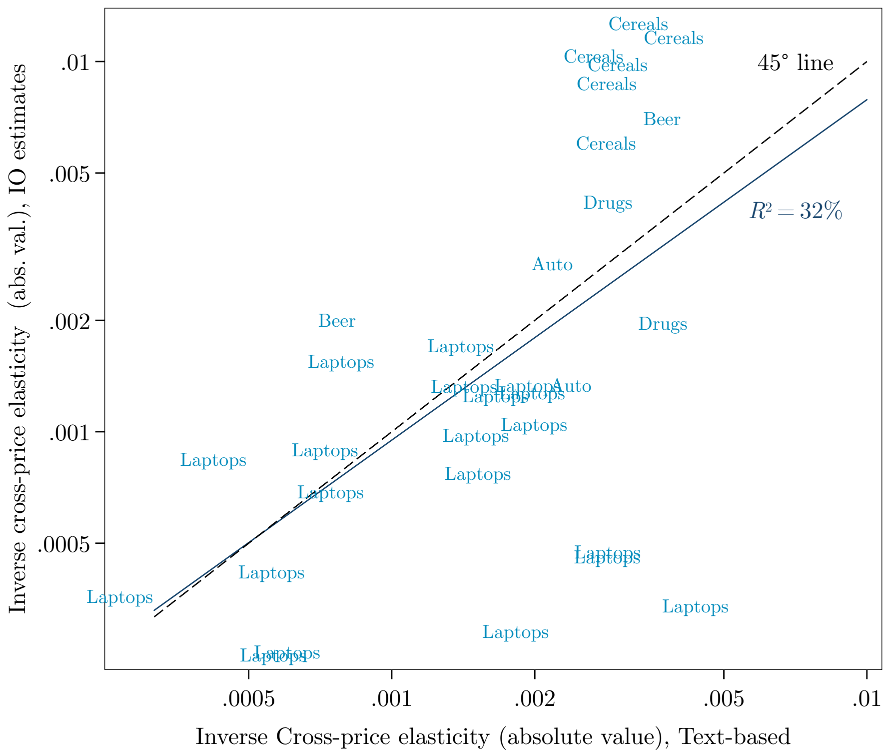
I look at non-targeted moments of the distribution of markups to evaluate the empirical performance of the model. The first non-targeted moment is the time-series correlation between the revenue-weighted average markup (μ_i) from my model and the corresponding estimate by DEU, outside of the calibration window (1983-2017). Figure 10 plots the two series. The navy line is the average markup based on my model. The light blue line is the series computed by DEU.

While only the average markup over 1957-1982 was targeted to calibrate $\tilde{\delta}$, the model-implied average markup tracks closely the time trend of the average markup of DEU after 1982.

Next, I examine how much cross-sectional variation in markups the model is able to generate. One major limitation of current general equilibrium models with market power is that the cross-sectional variation in markups that they can endogenously generate is nowhere close to that estimated by DEU.

In Figure 11, I plot the cross-section of markups obtained from the model against DEU's estimates. Both axes use a log scale. I include data from 1997, 2007 and 2017. As in DEU, I weight observations by revenue (weights are represented by circle sizes). On the one hand, it is fairly evident from the graph that there is more variation in the markups estimated by DEU than in the model-based estimates: the range is wider

Figure 12: Calibrated Elasticities vs. Microeconometric Estimates



Industry	Citation	Journal/Series
Auto	Berry, Levinsohn and Pakes (1995)	Econometrica
Beer	De Loecker and Scott (2016)	NBER Working Papers
Cereals	Nevo (2001)	Econometrica
Drugs	Chintagunta (2002)	Journal of Marketing Research
Laptops	Goeree (2008)	Econometrica

and the fitted regression slope (the dotted black line) is larger than one. On the other hand, I find that my model-based estimates capture a remarkable 77% of the cross-sectional variation in markups (as measured by the R-squared of a regression where the slope is constrained to one[21]).

Another important question from the point of view of model fit is how well my model-based estimates of the cross-price elasticity of demand correlate with the corresponding microeconometric estimates that were

[21] If we don't constrain the coefficient to one, the R^2 increases to 83%.

used to calibrate the parameter α. To this end, notice that equation (3.12) can be rearranged as:

$$\log \left| \frac{\partial \log p_i}{\partial \log q_j} \right| \;\approx\; \log \alpha + \log \left(\mathbf{a}_i' \mathbf{a}_j \, \frac{q_j}{p_i} \right) \tag{D.7}$$

where we use the symbol $\approx$ to denote the fact that measurement error contaminates what should be a linear-in-logs relationship.

By calibrating α, I was only able to choose the intercept of this linear relationship, not the slope nor the correlation. I will use the (untargeted) slope and the goodness-of-fit of equation (D.7) to evaluate the model fit.

In Figure 12, I plot the relationship between the microeconometric estimates (the left side of equation D.7) and the model-based estimates that use the calibrated value of α, together with the 45 degree line (the predicted slope) and a linear regression line. Each observation is a firm pair (i, j) and the labels reflect the industry study from whom the microeconometric estimates are sourced.

Despite the fact that we cannot affect the slope of the relationship by calibrating α, the slope of the linear fit between the two variables is statistically indistinguishable from one, and there is a strong positive correlation between the two series. The R^2 is 32%, which is particularly high if we consider that the microeconometric estimates were obtained by the respective authors using different assumptions about the underlying demand system, as well as different econometric methodologies.

E. Accounting for Fixed Costs and Intangible Capital

Figure 13: Replication of Figure 3, using Surplus net of Fixed Costs

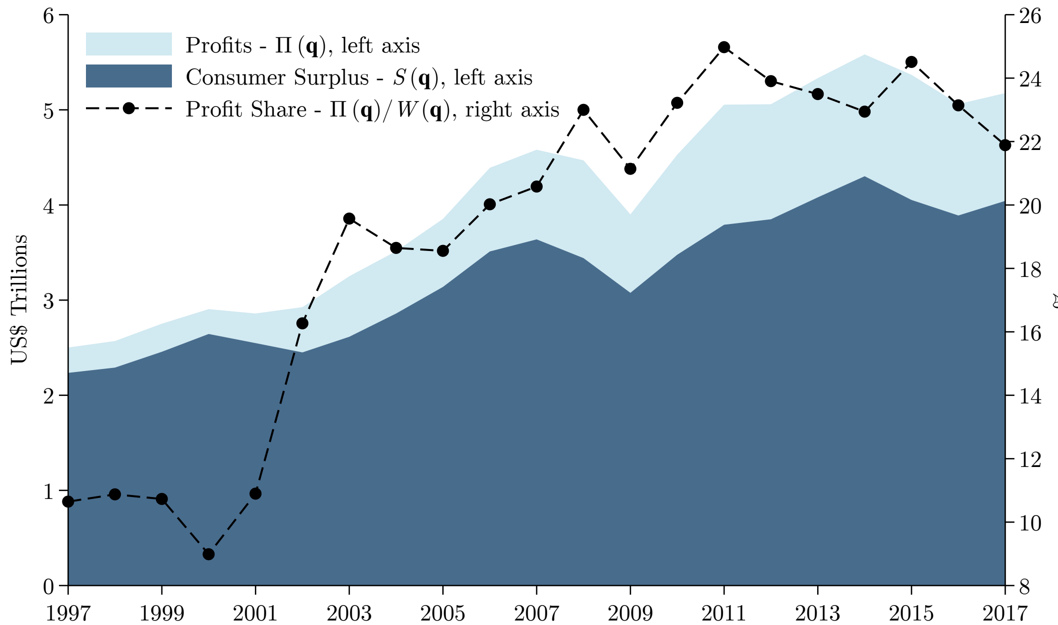
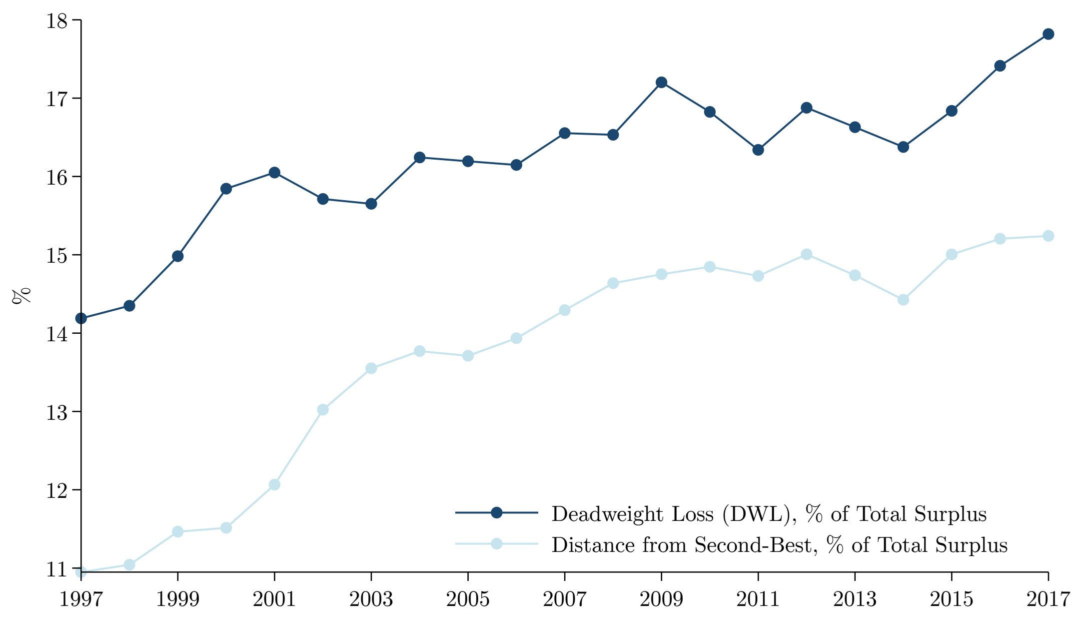


Figure 14: Replication of Figure 4, using Surplus net of Fixed Costs



In this appendix, I replicate Figures 3 and 4 using a different definition of aggregate profits, which is computed net of fixed costs. All results described in the main body of the paper are strengthened if we detract fixed costs. This suggests that fixed costs are not driving the measured increase in oligopoly power.

Figure 15: Replication of Figure 3, using Surplus net of Fixed Costs (Alternate Measure)

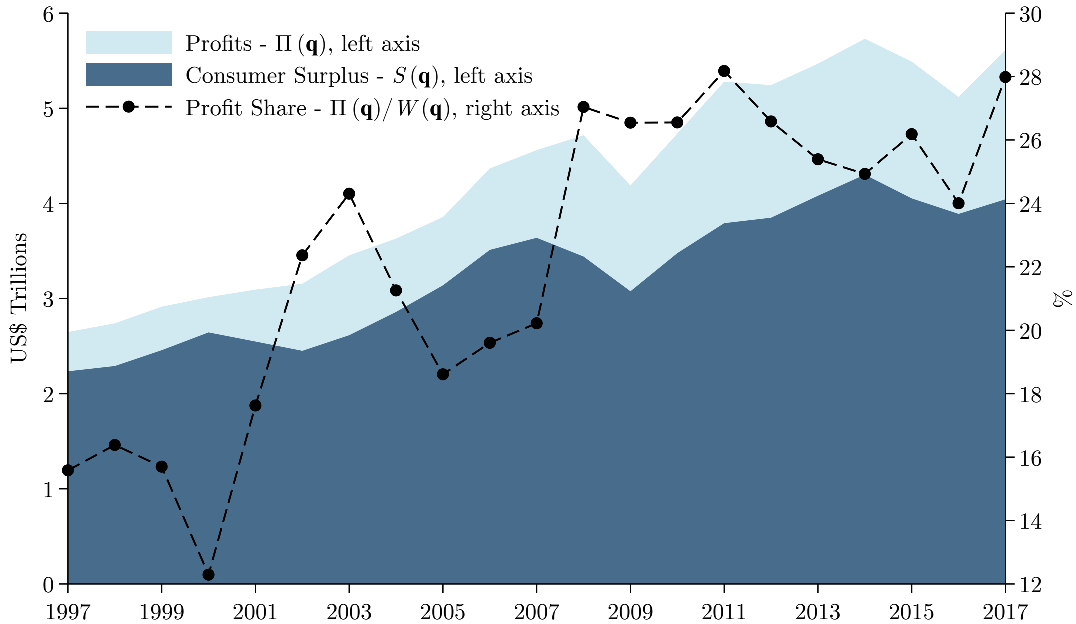
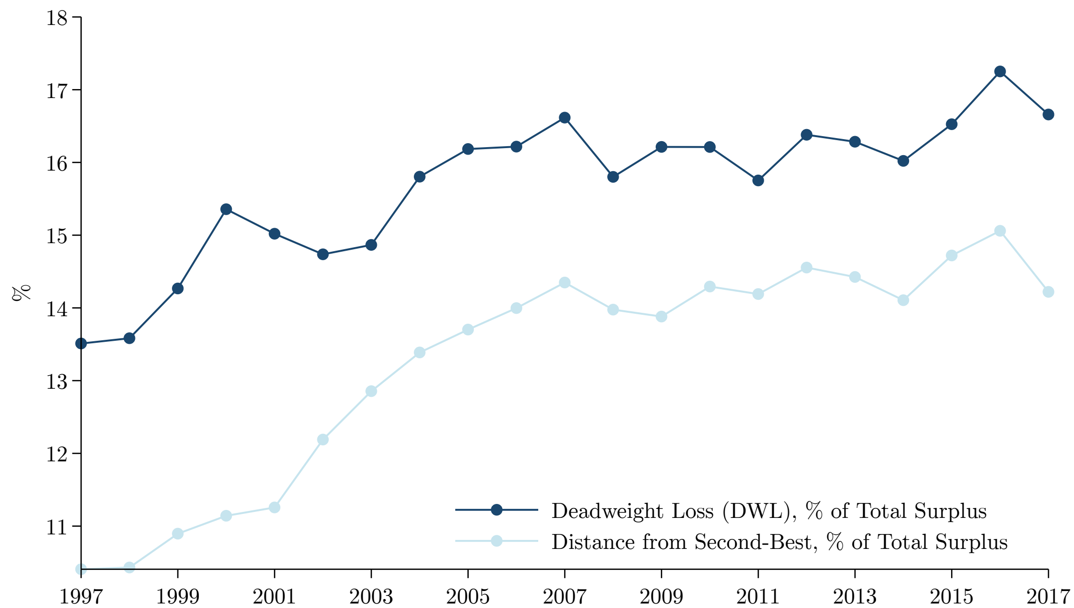


Figure 16: Replication of Figure 4, using Surplus net of Fixed Costs (Alternate Measure)



Figures 15 and 16 repeat this exercise using the alternate measure of fixed costs described in equation (5.2) (in which intangible investments are capitalized rather than expensed).

F. Bertrand Competition

In this Appendix I derive the equilibrium of the Bertrand counterpart of the Network Cournot game played by the oligopolists in the Model from Section 2. I focus on the special case of flat marginal cost – that is $\mathbf{\Delta} = \operatorname{diag}(\mathbf{0})$ (the case with non-flat marginal cost is less tractable). I start by writing the vector of economic profits in terms of the price vector $\mathbf{p}$:

$$\pi \quad = \quad \operatorname{diag}(\mathbf{p} - \mathbf{c})(\mathbf{I} + \mathbf{\Sigma})^{-1}(\mathbf{b} - \mathbf{p}) \tag{F.1}$$

Now let $\mathbb{D}$ and $\mathbb{O}$ be, respectively, the matrices containing the diagonal and off-diagonal elements of $(\mathbf{I} + \mathbf{\Sigma})^{-1}$ so that:

$$(\mathbf{I} + \mathbf{\Sigma})^{-1} \quad = \quad \mathbb{D} + \mathbb{O} \tag{F.2}$$

Then we can write:

$$\pi \quad = \quad \operatorname{diag}(\mathbf{p} - \mathbf{c})\left[\mathbb{D}(\mathbf{b} - \mathbf{p}) + \mathbb{O}(\mathbf{b} - \bar{\mathbf{p}})\right] \tag{F.3}$$

and take the first order condition firm-by-firm by taking the price vector of other firms (denoted by the upper bar as in Appendix A) as given:

$$0 \quad = \quad \mathbb{D}(\mathbf{b} - 2\mathbf{p}) + \mathbb{O}(\mathbf{b} - \mathbf{p}) - \mathbb{D}\,\mathbf{c} \tag{F.4}$$

which we can re-write in terms of $\mathbf{q}$ as:

$$0 \quad = \quad \mathbf{b} - \mathbf{c} - \left(\mathbb{D}^{-1} + \mathbf{I} + \mathbf{\Sigma}\right)\mathbf{q} \tag{F.5}$$

the corresponding Bertrand potential is

$$\Phi^B \quad = \quad \mathbf{q}'(\mathbf{b} - \mathbf{c}) - \frac{1}{2}\mathbf{q}'\left(\mathbf{I} + \mathbb{D}^{-1} + \Sigma\right)\mathbf{q} \tag{F.6}$$

and the Bertrand equilibrium is:

$$\mathbf{q}^B \quad = \quad \left(\mathbf{I} + \mathbb{D}^{-1} + \mathbf{\Sigma}\right)^{-1}(\mathbf{b} - \mathbf{c}) \tag{F.7}$$

Compare this with the Cournot Equilibrium (flat marginal cost):

$$\mathbf{q}^\Phi \quad = \quad (2\mathbf{I} + \mathbf{\Sigma})^{-1}(\mathbf{b} - \mathbf{c}) \tag{F.8}$$

and perfect competition:

$$\mathbf{q}^W \quad = \quad (\mathbf{I} + \mathbf{\Sigma})^{-1}(\mathbf{b} - \mathbf{c}) \tag{F.9}$$

because $\mathbb{D}^{-1}$ is a diagonal matrix whose diagonal components are between zero and one, we can see how Bertrand is a more "intense" form of competition than Cournot.

G. Further discussion of the GHL demand system

In this Appendix, I discuss more in depth how the GHL demand system introduced in this paper relates to other demand systems already encountered in the literature.

At a very basic level, GHL can be thought of as the hedonic, discrete counterpart of a type of preference aggregator that is already widely used in macroeconomics, trade, and industrial organization (see for example Asplund and Nocke 2006; Foster, Haltiwanger and Syverson 2008; Melitz and Ottaviano 2008):

$$U(\mathbf{q}) \;=\; q_0 + \alpha \int_{i \in \mathcal{I}} q_i \,\mathrm{d}i - \frac{\eta}{2} \left(\int_{i \in \mathcal{I}} q_i \,\mathrm{d}i \right)^2 - \frac{\gamma}{2} \int_{i \in \mathcal{I}} q_i^2 \,\mathrm{d}i \tag{G.1}$$

These linear-quadratic aggregators are frequently encountered in models with variable markups. My hedonic demand specification adds to this setup the following three features: 1) granular firms; 2) asymmetry in the degree of substitutability between different firm pairs; 3) the ability interpret the cross derivatives of the demand system in terms of product similarity.

Hedonic demand is something that I adopt from empirical industrial organization (see for example Berry, Levinsohn and Pakes, 1995). There are two main points of departure between GHL and standard IO demand models: the first is that the typical IO demand system is derived from a discrete choice model at the level of the individual consumer: it is obvious why, in a macroeconomics setting, with multiple products, it would be inappropriate to start from a discrete choice model (the consumer would have to choose *either* cars or laptops). The second difference is the functional form, which is linear instead of logit.

A key motivation for using linear demand as opposed to the widely-used CES aggregator is the ability of linear demand to produce heterogeneous markups. The reason why this is a desirable feature is that homogeneous markups, a feature of CES preferences, imply allocative efficiency (Dhingra and Morrow, 2019). While CES preferences have a number of desirable properties (such as the ability aggregate the output of firms with Pareto-distributed productivity), they become hard to rationalize within the context of this paper, not just empirically but also theoretically. If we try, for example, to write the isoelastic counterpart to the demand system in equation (2.21)

$$\log \mathbf{p} \;=\; \mathbf{b} - (\mathbf{I} + \mathbf{\Sigma}) \log \mathbf{q} \tag{G.2}$$

and derive the utility specification that generates it, we find that it will generally fail to satisfy the conditions for integrability, unless we impose the assumption that $\sigma_{ij} = 0 \;\forall\, ij$, which would make the model trivial. A more intuitive way to rephrase this mathematical fact is that in order to model the cross-price derivatives of the demand system in terms of product cosine similarities, we *need* a model with variable markups.

H. Additional Details on the "Constant IPO Ratio" Counterfactual

In this appendix, I provide additional details on how I implement the "constant IPO ratio" counterfactual. Without loss of generality, consider "spawning" $(N_1 - 1)$ additional firms that are identical to firm 1, except for the idiosyncratic characteristic. The profit and total surplus functions for each of these firms is:

$$\pi_1 \;=\; (b_1 - c_1)\, q_1 - \left(1 + \frac{\delta_1}{2}\right) q_1^2 - \alpha\,(Q_1 - q_1)\, q_1 - \alpha \sum_{j \neq 1} (\mathbf{a}_1' \mathbf{a}_j)\, q_1 q_j \tag{H.1}$$

$$w_1 \;=\; (b_1 - c_1)\, q_1 - \frac{1}{2}(1 + \delta_1)\, q_1^2 - \frac{\alpha}{2}(Q_1 - q_1)\, q_1 - \frac{\alpha}{2} \sum_{j \neq 1} (\mathbf{a}_1' \mathbf{a}_j)\, q_1 q_j \tag{H.2}$$

The first order condition is:

$$0 = \; b_1 - c_1 - (2 + \delta_1)\, q_1 - \alpha\,(Q_1 - q_1) - \alpha \sum_{j \neq 1} (\mathbf{a}_1' \mathbf{a}_j)\, q_j \tag{H.3}$$

where Q_1 is the joint output of all the N_1 resulting entities, and q_1 is the output of the individual firm. Because the child companies are all identical, in equilibrium they must produce the same quantity, therefore $q_1 = \frac{1}{N} Q_1$. We can then re-arrange:

$$0 = \; b_1 - c_1 - \frac{2 + \delta_1 + \alpha\,(N_1 - 1)}{N_1} Q_1 - \sum_{j \neq 1} \sigma_{1j} q_j \tag{H.4}$$

re-labelling Q_1 as q_1 we obtain this new set of first order conditions

$$0 \;=\; \mathbf{b} - \mathbf{c} - \begin{bmatrix} \frac{2+\delta_1+\alpha(N_1-1)}{N} q_1 \\ (2\mathbf{I} + \boldsymbol{\Delta}_2)\, \mathbf{q}_2 \end{bmatrix} - \boldsymbol{\Sigma}\mathbf{q} \tag{H.5}$$

where q_1 is no longer the output of the individual company of type 1 but the joint output of all the type-1 companies. The new Cournot equilibrium allocation maximizes the following modified potential:

$$\Phi(\mathbf{q}) \;=\; \mathbf{q}'(\mathbf{b} - \mathbf{c}) - \frac{1}{2}\mathbf{q}' \begin{bmatrix} \frac{2+\delta_1+\alpha(N-1)}{N} & 0 \\ 0 & (2\mathbf{I} + \boldsymbol{\Delta}_2) \end{bmatrix} \mathbf{q} - \frac{1}{2}\mathbf{q}'\boldsymbol{\Sigma}\,\mathbf{q} \tag{H.6}$$

The maximizer of $\Phi(\mathbf{q})$, which corresponds to the post-breakup equilibrium allocation, is

$$\mathbf{q}^{\Phi} = \left(\begin{bmatrix} \frac{2+\delta_1+\alpha(N-1)}{N} & 0 \\ 0 & (2\mathbf{I} + \boldsymbol{\Delta}_2) \end{bmatrix} \mathbf{q} + \boldsymbol{\Sigma} \right)^{-1} (\mathbf{b} - \mathbf{c}) \tag{H.7}$$

The total surplus and aggregate profit function have to be modified accordingly:

$$\Pi(\mathbf{q}) \;=\; \mathbf{q}'(\mathbf{b} - \mathbf{c}) - \mathbf{q}' \begin{bmatrix} \frac{1+\frac{1}{2}\delta_1+\alpha(N-1)}{N} & 0 \\ 0 & \mathbf{I} + \frac{1}{2}\boldsymbol{\Delta}_2 \end{bmatrix} \mathbf{q} - \mathbf{q}'\boldsymbol{\Sigma}\,\mathbf{q} \tag{H.8}$$

$$W(\mathbf{q}) \;=\; \mathbf{q}'(\mathbf{b} - \mathbf{c}) - \frac{1}{2}\mathbf{q}' \begin{bmatrix} \frac{1+\delta_1+\alpha(N-1)}{N} & 0 \\ 0 & (\mathbf{I} + \boldsymbol{\Delta}_2) \end{bmatrix} \mathbf{q} - \frac{1}{2}\mathbf{q}'\boldsymbol{\Sigma}\,\mathbf{q} \tag{H.9}$$

I. Differentiation and diversification: three examples

In this Appendix, I provide three simple numerical examples that illustrate how diversification could be associated, in the model, with decreasing, unchanged, or increasing competition. The direction of the effect clearly depends on whether cosine similarity increases, decreases or remains unchanged as a consequence of diversification.

I.1. Example 1: Similarity unaffected by diversification

Let us consider the example of two companies – say Apple and Samsung – that start by producing perfectly-substitutable computers (hence they compete fiercely) and then later diversify into producing mobile phones that are also perfectly-substitutable (this implies that there are no idiosyncratic characteristics, or $\alpha \to 0$). Quantitatively

$$\begin{array}{cccc} & \text{Apple} & \text{Samsung} & \\ & \downarrow & \downarrow & \\ \mathbf{A}_2 = & \left[\begin{array}{c} 1 \\ 0 \end{array}\right. & \left.\begin{array}{c} 1 \\ 0 \end{array}\right] & \begin{array}{l} \leftarrow \text{Computers} \\ \leftarrow \text{Mobile Phones} \end{array} \end{array} \tag{I.1}$$

after diversification occurs, this matrix becomes:

$$\tilde{\mathbf{A}}_2 = \begin{bmatrix} \sqrt{.5} & \sqrt{.5} \\ \sqrt{.5} & \sqrt{.5} \end{bmatrix} \tag{I.2}$$

It can be easily verified that, for any matrix $\mathbf{A}_1$ such that the columns of $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2$ have unit length (that is, regardless of how substitutable computers and mobile phones are):

$$\tilde{\mathbf{A}}_2'\mathbf{A}_1'\mathbf{A}_1\tilde{\mathbf{A}}_2 = \mathbf{A}_2'\mathbf{A}_1'\mathbf{A}_1\mathbf{A}_2 = \mathbf{A}'\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{I.3}$$

by computing the $(\mathbf{I}+\mathbf{\Sigma})^{-1}$ as $\alpha \to 0$ (no common characteristics) we find that the derivatives:

$$\lim_{\alpha\to 0}\left(-\frac{\partial q_i}{\partial p_i}\right) = \lim_{\alpha\to 0}\left(\frac{\partial q_i}{\partial p_{j\neq i}}\right) = \infty \tag{I.4}$$

hence, as long as Apple and Samsung diversify symmetrically (as measured by the change from $\mathbf{A}_2$ to $\tilde{\mathbf{A}}_2$) the model captures the fact that Apple and Samsung behave as perfect substitutes both before and after the change.

I.2. Example 2: Diversification reduces similarity

Next, consider a scenario in which only Samsung diversifies into mobile phones:

$$\tilde{\mathbf{A}}_2 = \begin{bmatrix} 1 & \sqrt{.5} \\ 0 & \sqrt{.5} \end{bmatrix} \tag{I.5}$$

and suppose, for simplicity, that $\mathbf{A}_1$ is equal to an identity matrix (hence there is no substitution between computers and phones). The derivative of q_{Samsung} with respect to p_{Apple} is 1.41, indicating that Apple and Samsung are no longer perfect substitutes from the point of view of the model.

This is not an accurate representation of the elasticities at the product level. However, we must be mindful that what the model takes as inputs are not product-level revenues and costs, but firm-level revenues and costs. When Apple raises computer prices, only the computer part of Samsung's product portfolio reacts to that price change.

Clearly, the ideal scenario would be to estimate the model at the product level. Lacking product-level data, I argue that it is appropriate to treat products Apple and Samsung as imperfect substitutes. If we treated Apple and Samsung as perfect substitutes after this diversification has occurred, we would be making the implicit assumption that Apple *computers* are perfect substitutes to Samsung *phones*.

In sum, in the example considered above, treating differentiated firms as imperfect substitutes is less-than-ideal. It is, however, still preferable to the alternative, which is to treat them as perfect substitutes while only being able to observe firm-level data.

I.3. Example 3: Diversification increases similarity

Finally, consider a scenario where, initially, Apple produces "mostly" computers and Samsung produces "mostly" phones. We continue to assume that all phones are perfect substitutes and that the same is true for computers. Also, phones do not interact strategically with computers. The two companies then become more diversified, with equal weight being applied to computers and phones:

$$\mathbf{A}_2 = \begin{bmatrix} \sqrt{8} & \sqrt{.2} \\ \sqrt{.2} & \sqrt{.8} \end{bmatrix} \qquad \tilde{\mathbf{A}}_2 = \begin{bmatrix} \sqrt{.5} & \sqrt{.5} \\ \sqrt{.5} & \sqrt{.5} \end{bmatrix} \tag{I.6}$$

In this case, the model initially sees the two companies as imperfect substitutes, although their products are, in each market, perfectly substitutable. After the product loadings are equalized, the cosine similarity between Apple and Samsung increases, and the two companies become eventually perfect substitutes in the model.

J. Labor Market Applications/Extensions

In this Appendix, I explore two labor-market applications of the linear-quadratic framework. The first is a model of skill worker monopoly: workers that have a unique set of skills (that are unlike other workers) are able to command higher wages. The second is a model of monopsony: firms that use inputs or labor markets not used by other firms are able to charge larger unit margins.

J.1. Worker Skill Monopoly

There are k types of workers. Each type is endowed with a unit of labor. There are n firms indexed by i that produce a homogeneous good (whose price is normalized to one) and act as price takers. They produce output q_i using a quadratic production function:

$$q_i = \mathbf{s}_i'\mathbf{b}_i - \frac{1}{2}\mathbf{s}_i'\mathbf{s}_i \tag{J.1}$$

where $\mathbf{s}_i$ is the vector of skill units that depend on the vector of labor units $\mathbf{h}_i$ hired:

$$\mathbf{s}_i = \mathbf{A}\mathbf{h}_i \tag{J.2}$$

As in the model from Section 2, the matrix $\mathbf{A}$ transforms labor units into skill units, and $\|\mathbf{a}_i\| = 1$. Hence

$$q_i = \mathbf{h}_i'\mathbf{A}'\mathbf{b}_i - \frac{1}{2}\mathbf{h}_i'\mathbf{A}'\mathbf{A}\mathbf{h}_i \tag{J.3}$$

The optimality condition for firm i is that the marginal revenue product of each type of labor is equal to the corresponding wage

$$\mathbf{MRPL}_i = \mathbf{A}'\mathbf{b}_i - \mathbf{A}'\mathbf{A}\mathbf{h}_i = \mathbf{w} \tag{J.4}$$

The labor demand for firm i is:

$$\mathbf{h}_i = \left(\mathbf{A}'\mathbf{A}\right)^{-1}\left(\mathbf{A}'\mathbf{b}_i - \mathbf{w}\right) \tag{J.5}$$

Summing across firms and defining $\mathbf{b} = \sum_i \mathbf{A}'\mathbf{b}_i$

$$1 = \left(\mathbf{A}'\mathbf{A}\right)^{-1}\left(\mathbf{b} - n\mathbf{w}\right) \tag{J.6}$$

Hence j's equilibrium wage is equal to the marginal revenue product of the first hour (b_j) less the average cosine similarity of j to every other worker:

$$w_j = b_j - \frac{1}{n}\sum_{m=1}^{k}\mathbf{a}_j'\mathbf{a}_m \tag{J.7}$$

J.2. Monopsony

There are k types of workers and n firms indexed by i that act as price-takers in goods markets and face output price vector $\mathbf{p}$. The firms produce using a Leontief production function, so that the following vector of labor inputs $\mathbf{h}$ is required in order to produce output vector $\mathbf{q}$:

$$\mathbf{h} = \mathbf{T}\mathbf{q} \tag{J.8}$$

Each dimension of $\mathbf{h}$ is a worker and each dimension of $\mathbf{q}$ is a firm. Once again, we normalize $\|\mathbf{t}_i\| = 1$. Assume that the labor supply function is linear, so that i's wage is:

$$w_i = \lambda_i h_i \tag{J.9}$$

The total cost of firm i is equal to

$$\mathrm{TC}_i = \sum_j w_j x_{ij} = q_i \sum_j w_j t_{ji} = \lambda_i q_i \sum_j \sum_i q_j t_{ij} t_{ji} \tag{J.10}$$

We can then write the vector of profit as:

$$\pi = \mathrm{diag}\,(\mathbf{q})\,\mathbf{p} - \boldsymbol{\Lambda} \cdot \mathrm{diag}\,(\mathbf{q})\,\mathbf{T}'\mathbf{T}\mathbf{q} \tag{J.11}$$

where

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{J.12}$$

Firms compete on quantity, taking the output of all other firms as given. The potential function is:

$$\Phi\,(\mathbf{q}) = \mathbf{q}'\mathbf{p} - \boldsymbol{\Lambda} \cdot \mathbf{q}\,(\mathbf{I} + \mathbf{T}'\mathbf{T})\,\mathbf{q} \tag{J.13}$$

Taking the first order condition we find the Nash equilibrium. It is once again a measure of centrality - this time in the input/technological space:

$$\mathbf{q}^{\Phi} = (\mathbf{I} + \boldsymbol{\Lambda}\mathbf{T}'\mathbf{T})^{-1}\,\mathbf{p} \tag{J.14}$$

Because q_i is proportional to the unit margin, one implication of the equation above is that firms that use labor not used by other firms have monopsony power in those labor markets.

K. Proofs and Derivations

Proof to Lemma 1. To prove the lemma we use the fact that, at the Cournot equilibrium, the following two relationships hold:

$$\pi_i \;=\; \left(1+\frac{\delta_i}{2}\right)q_i^2 \qquad \text{and} \qquad s_i \;=\; \frac{1}{2}\left(q_i^2 + \sum_{j\neq i}\sigma_{ij}q_i q_j\right) \tag{K.1}$$

then:

$$\frac{\pi_i}{s_i} \;=\; \frac{\left(1+\frac{\delta_i}{2}\right)\cdot q_i^2}{\frac{1}{2}\cdot q_i\left(q_i + \sum_{j\neq i}\sigma_{ij}q_j\right)} \;=\; (2+\delta_i)\cdot\frac{q_i}{q_i + \sum_{j\neq i}\sigma_{ij}q_i} \;\equiv\; (2+\delta_i)\cdot\omega_i \tag{K.2}$$

□

Proof to Lemma 2. To simulate a merger (or a collusion), we sum the first rows of the profit function that correspond to the merging firms:

$$\begin{bmatrix}\Pi_1\\ \pi_2\end{bmatrix} = \begin{bmatrix}\mathbf{q}_1'\\ \mathbf{Q}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}-\mathbf{c}_1\\ \mathbf{b}-\mathbf{c}_2\end{bmatrix} - \begin{bmatrix}\mathbf{q}_1\\ \mathbf{Q}_2\end{bmatrix}'\begin{bmatrix}\left(\mathbf{I}+\frac{1}{2}\boldsymbol{\Delta}_1\right)\mathbf{q}_1\\ \left(\mathbf{I}+\frac{1}{2}\boldsymbol{\Delta}_2\right)\mathbf{q}_2\end{bmatrix} - \begin{bmatrix}\mathbf{q}_1\\ \mathbf{Q}_2\end{bmatrix}'\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}\begin{bmatrix}\mathbf{q}_1\\ \mathbf{q}_2\end{bmatrix} \tag{K.3}$$

where

$$\mathbf{Q}_2 \overset{\text{def}}{=} \operatorname{diag}(\mathbf{q}_2) \tag{K.4}$$

I have partitioned the profits vector into a scalar Π_1, which collects the joint profits of the new entity, and vector π_2, in which I stack the profits of all the other companies that are not included in the merger. If there are n firms and two of them are merging, this is a $(n-1)$ dimensional column vector. The system of first order condition solved by the surviving firms is:

$$\begin{bmatrix}0\\ 0\end{bmatrix} = \begin{bmatrix}\mathbf{b}-\mathbf{c}_1\\ \mathbf{b}-\mathbf{c}_2\end{bmatrix} - \begin{bmatrix}(2+\delta_1)\,\mathbf{q}_1\\ (2\mathbf{I}+\boldsymbol{\Delta}_2)\,\mathbf{q}_2\end{bmatrix} - \begin{bmatrix}2\cdot\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}\begin{bmatrix}\mathbf{q}_1\\ \bar{\mathbf{q}}_2\end{bmatrix} \tag{K.5}$$

□

Proof to Proposition 3. An atomistic firm z has the following cost function:

$$h(z) \;=\; \frac{q^2(z)}{2z} \tag{K.6}$$

Let p_{n+1} be the price of the good sold by the atomistic firms (recall we assumed they produce a homogeneous product). Because firms behave competitively and price at marginal cost, it must be the case that:

$$q(z) \;=\; p_{n+1}\cdot z \tag{K.7}$$

hence labor supply and profits are given by:

$$h(z) = \frac{p_{n+1}^2}{2} \cdot z; \qquad \pi(z) = \frac{p_{n+1}^2}{2} \cdot z \tag{K.8}$$

Because the atomistic firms pay an entry cost of one unit of labor, the productivity cutoff for entry, which we call $z_{\min}$, is given by:

$$1 = \frac{p_{n+1}^2}{2} \cdot z_{\min} \tag{K.9}$$

Let us now compute aggregate output and aggregate labor

$$\begin{aligned} q_{n+1} &= \int_{z_{\min}}^{\infty} q(z)\,\mathrm{d}F(z) &&= p_{n+1}\int_{z_{\min}}^{\infty} z \cdot f(z)\,\mathrm{d}z &&= \sqrt{\tfrac{2}{z_{\min}}}\int_{z_{\min}}^{\infty} \tfrac{\beta-1}{z^{\beta}}\,\mathrm{d}z &&= \sqrt{2}z_{\min}^{\frac{1}{2}-\beta} \\ h_{n+1} &= \int_{z_{\min}}^{\infty} h(z)\,\mathrm{d}F(z) &&= \tfrac{p_{n+1}^2}{4}\int_{z_{\min}}^{\infty} z \cdot f(z)\,\mathrm{d}z &&= \tfrac{1}{z_{\min}}\int_{z_{\min}}^{\infty} \tfrac{\beta-1}{z^{\beta}}\,\mathrm{d}z &&= z_{\min}^{-\beta} \end{aligned} \tag{K.10}$$

By writing the productivity cutoff $z_{\min}$ in terms of aggregate output q_{n+1} and plugging it in the expression for aggregate cost h_{n+1}, we find the aggregate cost function:

$$h_{n+1} = \left[\left(\frac{q_{n+1}}{\sqrt{2}}\right)^{\frac{1}{\frac{1}{2}-\beta}}\right]^{(-\beta)} = \left(\frac{q_{n+1}}{\sqrt{2}}\right)^{\frac{2\beta}{2\beta-1}} \tag{K.11}$$

by taking the limit $\beta \to 1^+$, we can see that the expression above converges to the quadratic form:

$$h_{n+1} = \frac{q_{n+1}^2}{2} \tag{K.12}$$

□

Proof to Lemma 4. The markup is equal to:

$$\mu_i = \frac{p_i}{\mathrm{MC}_i} = \frac{p_i}{c_i + \delta q_i} = \frac{p_i q_i}{c_i q_i + \delta q_i^2} = \frac{p_i q_i}{h_i + \frac{\delta}{2} q_i^2} \tag{K.13}$$

we can then use the fact that $\pi_i = (1 + \delta_i/2)\, q_i^2 = (p_i q_i - h_i)$ to write:

$$\mu_i = \frac{p_i q_i}{h_i + \frac{\delta}{2(1+\delta_i/2)}\pi_i} = \frac{(2+\delta_i)\, p_i q_i}{(2+\delta_i)\, h_i + \delta_i \pi_i} = \frac{(2+\delta_i)\, p_i q_i}{2 \cdot h_i + \delta_i p_i q_i} \tag{K.14}$$

□
